MAT 201X: Applied Analysis

Greg DePaul

September 17, 2020

1 Metric and Normed Spaces

1.1 Metrics

Given a set X of objects, we aim to study the relationship between these objects. To do so, we have the ingeneous idea to put a distance function to measure the betweenness of these objects.

Definition 1.1. *Let X be a set. A functions $d : X \times X \to \mathbb{R}$ is a* metric *on X if it satisfies the following conditions:*

- $d(x, y) \geq 0$ *for all* $x, y \in X$ *and* $d(x, y) = 0$ *if and only if* $x = y$.
- $d(x, y) = d(y, x)$
- $d(x, z) \leq d(x, y) + d(y, z)$

Note 1.2. *The positivity can be derived from the other conditions:*

$$2d(x, y) = d(x, y) + d(y, x) \geq d(x, x) = 0$$

Example 1.3. $(\mathbb{R}, |\cdot|)$ *is a metric space.*

Example 1.4. $(\mathbb{C}, |\cdot|)$ *is a metric space.*

Example 1.5. *If X is a nonempty set, we may define*

$$d(x, y) = \begin{cases} 0 & x = y \\ 1 & x \neq y \end{cases}$$

Then we have a metric space on (X, d)

Example 1.6. *Given (X, d), we can construct a new metric space by modifying X or d.*

1. *We can modify d in several ways:*

$$d'(x, y) = \lambda d(x, y)$$
$$d''(x, y) = \frac{d(x, y)}{1 + d(x, y)}$$
$$d'''(x, y) = \ln(1 + d(x, y))$$
$$d''''(x, y) = \min\{1, d(x, y)\}$$

In any of the above examples, we can identifying a function that which is composed by the original metric function d.

Exercise 1.7. *On which conditons of f will $f \circ d$ become a metrix on X? Homework!*

2. *We can modify X in several ways:*
   - *If $Y \subset X$, then the restriction $d|_Y$ is a metric on Y, resulting in the space $(Y, d|_Y)$*
   - *But what of the case of $\mathbb{S}^1 \subset \mathbb{R}^2$? The distances by restriction will be the lengths of the secant, not the arc. But, we often want the distance to be intrinsic to the space, hence this merits a metric that gives the length of the shorter arc between two points.*

Example 1.8. *Suppose (X, d_X) and (Y, d_Y) are two metric spaces. Then we may define a metric d_{XY} on the product space $X \times Y = \{(x, y) : x \in X, y \in Y\}$ by*

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

which leads to the space $(X \times Y, d_{X\times Y})$.
We may also define

$$d'_{X\times Y}((x_1, y_1), (x_2, y_2)) = \sqrt{d_X(x_1, x_2)^2 + d_Y(y_1, y_2)^2}$$

In general, given metric spaces $\{(X_i, d_i)\}_{i=1}^n$ and any $p \geq 1$, we may define on the space

$$X_1 \times X_2 \times \cdot \times X_n$$

and equip with the metric

$$d_p((x_1, \cdots, x_n), (y_1, \cdots, y_n)) = \left[\sum_{i=1}^{n} d_i(x_i, y_i)^p\right]^{\frac{1}{p}}$$

to construct a new space.

Example 1.9. *Returning to $(\mathbb{R}, |\cdot|)$, we define $\mathbb{R}^n = \mathbb{R} \times \cdots \mathbb{R}$ and define the metric*

$$d_p((x_1, \cdots, x_n), (y_1, \cdots, y_n)) = \left[\sum_{i=1}^{n} |x_i - y_i|^p\right]^{\frac{1}{p}}$$

For the case $p = 2$, we refer to this metric as the standard euclidean metric.

Example 1.10. *Let $A = \{a_i\}_{i=1}^k$ be a k-character alphabet. We equip this with the discrete metric. Then*

$$A^n = \{(x_1, \ldots, x_n) : x_i \in A\}$$

is the set of n-letter words. A metric on such a set could be:

$$d_1(x, y) = \sum_{i=1}^{n} d(x_i, y_i) := \#\{i : x \neq y\} \text{ is the Word Metric}$$

$$\text{i.e. } d_1(sun, son) = 1$$

$$\text{i.e. } d_1(cat, dog) = 3$$

1.2 Norms

Definition 1.11. *Let X be a vector space over $\mathbb{R}$ or $\mathbb{C}$. Then a norm on X, denoted by $\|\cdot\|$ if a function $\|\cdot\| : X \to [0, \infty)$ with the following properties*

- $\|v\| \geq 0$ *and* $\|v\| = 0$ *if and only if* $v = 0$
- $\|\lambda v\| = |\lambda| \, \|v\|$ *with* $\lambda \in \mathbb{R}$ *or* $\mathbb{C}$.

- $\|v+w\| \leq \|v\| + \|w\|$

Note 1.12. *A normed space induces a metric space: Let $(X, \|\cdot\|)$ be a normed linear space. Then*

$$d(x,y) = \|x-y\|$$

defines a metric on X.

Question 1.13. *Suppose d a metric on a vector space. Can d always be induced from a norm? No! Homework!*

Definition 1.14. *A subset A of a vector space X is* <u>*convex*</u> *if for all $x, y \in A$*

$$tx + (1-t)y \in A$$

for all $t \in [0,1]$

Proposition 1.15. *The closed unit ball of a normed linear space $(X, \|\cdot\|)$, $\bar{B}_1(0) = \{x \in X : \|x\| \leq 1\}$ is convex.*

Proof. Suppose $\|\cdot\|$ is a norm on X. Then for all $x, y \in \bar{B}_1(0)$ and any $t \in [0,1]$

$$\|tx + (1-t)y\| \leq t\|x\| + (1-t)\|y\| \leq t + (1-t) = 1$$

∎

Example 1.16. *The Euclidean space $\mathbb{R}^n$ is a normed linear space whose norm is expressed by*

$$\|x\| = \sqrt{x_1^2 + \ldots + x_n^2}$$

Example 1.17. *There are other natural norms on $\mathbb{R}^n$ such as the 1-norm*

$$\|x\|_1 = |x_1| + |x_2| + \ldots + |x_n|$$

The p-norm where $p \geq 1$

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}$$

The ∞-norm

$$\|x\|_\infty = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$$

Exercise 1.18. *If $\|\cdot\|_p$ a norm for $0 < p < 1$? No! It fails the triangle inequality. You can also consider it's closed ball, which, by visual inspection, is concave.*

1.3 Convergence of Sequences

Let (X, d) be a metric space.

Definition 1.19. *A* <u>*sequence*</u> *in X is a function $X : \mathbb{N} \to X$, commonly denoted by $\{x_n\}_{n=1}^\infty$.*

Definition 1.20. *Let $\{x_n\}_{n=1}^\infty$ be sequence in X and $x \in X$. Then $\{x_n\}_{n=1}^\infty$* <u>*converges*</u> *to x provided for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$*

$$d(x, x_n) < \epsilon$$

Definition 1.21. *A sequence in a metric space X is a* <u>*Cauchy*</u> *if for every $\epsilon > 0$, there exists N such that*

$$d(x_n, x_m) < \epsilon$$

for all $n, m > N$.

Proposition 1.22. *If $\{x_n\}$ converges, then $\{x_n\}$ is Cauchy.*

Note 1.23. *Unfortunately the converse is not true. Consider the space $((0,1), |\cdot|)$ and the sequence $\{\frac{1}{n}\}$ is Cauchy by not convergent in $((0,1), |\cdot|)$.*

Proposition 1.24. *If $\{x_n\}$ is Cauchy, then $\{x_n\}$ is bounded.*

Proposition 1.25. *Suppose $\{x_n\}$ is Cauchy. If it has a convergent subsequence $\{x_{n_k}\}$, then $\{x_n\}$ is convergent.*

Definition 1.26. *A metric space (X, d) is called complete if every Cauchy sequence in (X, d) converges in (X, d).*

Example 1.27. $(\mathbb{R}, |\cdot|)$ *is a complete normed space.*

Example 1.28. $(\mathbb{Q}, |\cdot|)$ *is* ***not*** *a complete metric space.*

Example 1.29. *Let d be the discrete metric on X. Then (X, d) is a complete metric space.*

1.4 Open and Closed Sets

Let (X, d) be a metric space.

Definition 1.30. *For all $x \in X, r > 0$, we define the open ball*

$$B_r(x) := \{y \in X : d(y, x) < r\}$$

Definition 1.31. *For all $x \in X, r > 0$, we define the closed ball*

$$\bar{B}_r(x) := \{y \in X : d(y, x) \leq r\}$$

Example 1.32. *Let d be the discrete metric on X. Then*

$$B_r(x) = \begin{cases} X & r > 1 \\ \{x\} & r \leq 1 \end{cases} \qquad \bar{B}_r(x) = \begin{cases} X & r \geq 1 \\ \{x\} & r < 1 \end{cases}$$

Definition 1.33. *A set $U \subset X$ is open provided for every $x \in U$, we can identify a $r > 0$ such that*

$$B_r(x) \subset U$$

Definition 1.34. *A set $F \subset X$ is closed provided $F^c = X \setminus F$ is open.*

Proposition 1.35. *$F \subset X$ is closed if and only if every convergent sequence of elements in F converges to a limit in F.*

Proof. Homework. ∎

Proposition 1.36. *Let (X, d) be a metric space. Then*

- *$\emptyset, X$ are both closed and open.*
- *A finite intersection of open sets is open.*
- *A finite union of closed sets is closed.*
- *Arbitrary unions of open sets is still open.*
- *Arbitrary intersections of closed sets is still closed.*

Note 1.37. *Arbitrary intersections of open sets is not necessarily open.*

Proposition 1.38. *Let (X,d) be a metric space and $F \subset X$. Then*

1. *If F is complete, then F is closed.*
2. *If X is complete, and $F \in X$ is closed, then F is complete.*

So if X is complete, then ($F \subset X$ is closed if and only if F is complete).

Definition 1.39. *For any $A \subset X$, the closure of A is the smallest closed subset of X that contains A, denoted by $\bar{A}$. Moreover,*

$$\bar{A} := \bigcap_{F \text{ closed } \subset X; A \subset F} F$$

Equivalently,

$$\bar{A} := \{x \in X : \text{ there exists a sequence in } X \text{ that converges to } x\}$$

Definition 1.40. *A set $A \subset X$ is dense if $\bar{A} = X$. In other words, for every point $x \in X$, there exists $\{a_n \in A\}_{n=1}^{\infty}$ such that*

$$a_n \to x$$

Example 1.41. *Let $X = C([0,1])$ and let $d(f,g) = \|f-g\|_\infty$. Then the set of all polynomials*

$$\{p : [0,1] \to \mathbb{R} : p \text{ is a polynomial }\}$$

is dense in (X,d).

Proposition 1.42. *A set $A \subset X$ is dense if and only if for every $\epsilon > 0$, and every $x \in X$, $B_\epsilon(x) \cap A \neq 0$.*

Definition 1.43. *A metric space (X,d) is seperable if it has a countable dense subset.*

1.5 Continuity

Let (X, d_X) and (Y, d_Y) be a metric space.

Definition 1.44. *Let $x_0 \in X$ and $f : X \to Y$ be a function. Then f is continuous at x_0 if for all $\epsilon > 0$, $\exists \delta > 0$ such that*

$$d_X(x, x_0) < \delta \implies d_Y(f(x), f(x_0)) < \epsilon$$

That is,

$$f(B_\delta(x_0)) \subset B_\epsilon(f(x_0))$$

Further, f is continuous on a subset $E \subset X$ provided f is continuous at every point $x \in E$.

Note 1.45. *$f : X \to Y$ is continuous on $E \subset X$, then the restriction*

$$f|_E : E \to Y$$

is continuous (a priori). The converse is not true.

Example 1.46. *Define the Dirichlet function*

$$f(x) = \mathcal{X}_{\mathbb{Q}}(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{Q}^c \end{cases}$$

Then $E = \mathbb{Q}$ serves as a example of the note above.

Theorem 1.47. *Suppose $f : X \to Y$ is a function between metric space. Then the following statements are equivalent:*

1. *f is continuous on X*
2. *If $U \subset Y$ is open, then $f^{-1}(U) = \{x \in X : f(x) \in U\}$ is open in X*
3. *If $\lim_{n\to\infty} x_n = x$ in X, then $\lim_{n\to\infty} f(x_n) = f(x)$*

Note 1.48. *(2) is the topological definition. (3) is suitable for metric spaces.*

Proof.

- (1) $\implies$ (2)
  Suppose $f : X \to Y$ is continuous. Then for all $U \subset Y$ open, and let $x_0 \in f^{-1}(U)$. Then $f(x_0) \in U$, which is open in Y. So there exists $\epsilon > 0$ such that
  $$B_\epsilon(f(x_0)) \subset U$$
  Since f is continuous, we can identify $\delta > 0$ such that
  $$f(B_\delta(x_0)) \subset B_\epsilon(f(x_0)) \subset U \implies B_\delta(x_0) \subset f^{-1}(U)$$
  Thus, $f^{-1}(U)$ is open.
- (2) $\implies$ (3)
  Let $\{x_n\}$ be a sequence in X and $\lim_{n\to\infty} x_n = x$. We want to show that $\lim_{n\to\infty} f(x_n) = f(x)$. Recall for all $\epsilon < 0$, since $B_\epsilon(f(x))$ is open in Y, $f^{-1}(B_\epsilon(f(x))$ is open in X. Since $x \in f^{-1}(B_\epsilon(f(x))$, then $\exists \delta > 0$ such that
  $$B_\delta(x) \subset f^{-1}(B_\epsilon(f(x))$$
  That is,
  $$f(B_\delta(x)) \subset B_\epsilon(f(x))$$
  Now since $x_n \to x$, there exists N such that for all $n > N$
  $$d_X(x_n, x) < \delta \implies x_n \in f(x_n) \in B_\epsilon(f(x)) \iff d_Y(f(x), f(x_n)) < \epsilon$$
  Therefore, $f(x_n) \to f(x)$
- (3) $\implies$ (1)
  Let $x_0 \in X$. Suppose that $x_n \to x$ implies $f(x_n) \to f(x)$. We want to show that f is continuous at x_0. Assume not. Then $\exists \epsilon_0 > 0$ such that for any $\delta > 0$
  $$f(B_\delta(x_0)) \not\subset B_{\epsilon_0}(f(x_0))$$
  In particular, for all $k \in \mathbb{N}$, let $\delta = \frac{1}{k}$. Then
  $$f(B_{\frac{1}{k}}(x_0)) \not\subset B_{\epsilon_0}(f(x_0))$$
  Therefore, $\exists x_k \in f(B_{\frac{1}{k}}(x_0))$ such that $f(x_k) \notin B_{\epsilon_0}(f(x_0))$. So $x_k \to x$ but $d_Y(f(x_k), f(x)) \geq \epsilon$. Contradiction!

∎

Definition 1.49. *A function $f : X \to Y$ is uniformly continuous if for all $\epsilon > 0$, there exists $\delta(\epsilon) > 0$ such that*
$$d_Y(f(x), f(y)) < \epsilon$$
whenever $d_X(x, y) < \delta$.

Definition 1.50. *A function $f : X \to Y$ is upper semi-continuous if for every sequence $\{x_n\} \to x$, we have that*

$$\limsup_{n\to\infty} f(x_n) \leq f(x)$$

Similarly, $f : X \to Y$ is lower semi-continuous if for every sequence $\{x_n\} \to x$, we have that

$$\liminf_{n\to\infty} f(x_n) \geq f(x)$$

Note 1.51. *f is continuous if and only if f is both upper and lower semi-continuous.*

Proposition 1.52. *If f lower semi-continuous on $[0, 1]$, then f achieves its minimum value on $[0, 1]$.*

Proof. Taking any minimizing sequence $\{x_n\}$

$$\lim_n x_n = \inf_{x\in[0,1]} f(x)$$

$\{x_n\}$ has a convergent subsequence. $x_{n_k} \to x^*$. f is lower semi-continuous

$$\implies f(x^*) \leq \liminf f(x_{n_k}) = \inf_{x\in[0,1]} f(x) \implies f \text{ is minimized at } x_0$$

■

1.6 The Completion of a Metric Space

Definition 1.53. *Let (X, d) and $(\tilde{X}, \tilde{d})$ be metric spaces. Then an isometric embedding is a one-to-one map $i : X \to \tilde{X}$ such that for all $x, y \in X$*

$$\tilde{d}(i(x), i(y)) = d(x, y)$$

Definition 1.54. *An isometric isomorphism is a surjective isometric embedding. Moreover, Two Metric Spaces (X, d) and (Y, d') are isometric if there exists a bijection $h : X \to Y$ that preserves the metric.*

Example 1.55. *Folds of a piece of paper should maintain the same topology intrinsic to that sheet of paper. This is desirable so that we can describe the same object as it changes.*

Remark 1.56. *Two isometric spaces are indistinguishable as metric spaces. They share all properties that can be expressed completely in terms of distances.*

Definition 1.57. *$(\tilde{X}, \tilde{d})$ is the completion of (X, d) isometric isomorphism if*

1. *$\exists i : X \to \tilde{X}$ an isometric embedding.*
2. *$i(X)$ is dense in $\tilde{X}$*
3. *$\tilde{X}$ is complete*

Theorem 1.58. *For every metric space (X, d) there exists a completion $(\tilde{X}, \tilde{d})$. Moreover, this completion is unique up to the isometric isomorphism.*

Example 1.59. *$\mathbb{Q}$ is incomplete. It's completion is $\mathbb{R} = \mathbb{Q} \cup \mathbb{Q}^c$. How do you construct this? Consider*

$$\pi = 3, 3.1, 3.14, 3.141, \ldots$$

So this number $\pi \notin \mathbb{Q}$ can be represented by a cauchy sequence $\{q_n\}_n \subset \mathbb{Q}$. But, what if two cauchy sequences converge to π? How do we reconcile this?

Using the example as motivation, we return to the proof.

Proof. • **Step 1**: Construct $(\tilde{X}, \tilde{d})$ using the Cauchy sequences in (X, d). Define

$$\mathcal{C} := \text{ the set of all Cauchy sequences in } (X, d)$$

Then for all $x, y \in \mathcal{C}$, we can write $x = \{x_n\}, y = \{y_n\}$. We define the metric $\tilde{d}$ by

$$\hat{d}(x, y) := \lim_{n \to \infty} d(x_n, y_n)$$

Claim: $d(x_n, y_n)$ is a Cauchy sequence in $\mathbb{R}$.

$$|d(x_n, y_n) - d(x_m, y_m)| \leq |d(x_n, y_n) - d(x_n, y_m)| + |d(x_n, y_m) - d(x_m, y_m)| \leq d(y_n, y_m) + d(x_n, x_m) < 2\epsilon \to 0$$

By the completeness of $(\mathbb{R}, |\cdot|)$, the limit $\lim_{n\to\infty} d(x_n, y_n)$ exists, showing $\hat{d}(x, y)$ is well-defined. BUT, we need to show that $\hat{d}$ is a metric.

- We inherit symmetry, positivity, and the triangle inequality from the sheer definition of our distance $\tilde{d}$ being defined in terms of d.
- However, we see that

$$\hat{d}(x, y) = 0 \not\Longrightarrow x = y$$

  This is called a pseudo-metric on $\mathcal{C}$. So we need to change the space we are working on to fix this! Define for all $x, y \in \mathcal{C}$

$$x \sim y \iff \hat{d}(x, y) = 0$$

  Then $\sim$ is an equivalence class! We now define

$$\tilde{X} := \mathcal{C}/\sim := \{[x] : x = \{x_n\}_n \in \mathcal{C}\}$$

So we edit the definition of $\tilde{d}$ slightly to be

$$\tilde{d}([x], [y]) := \hat{d}(x, y)$$

Claim: This distance function is well-defined for any representation of $[x]$ or $[y]$
Suppose that

$$x = \{x_n\} \in [x]; x' = \{x'_n\} \in [x]$$
$$y = \{y_n\} \in [y]; y' = \{y'_n\} \in [y]$$

Observe:

$$|d(x_n, y_n) - d(x'_n, y'_n)| \leq d(x_n, x'_n) + d(y_n, y'_n) < 2\epsilon \to 0$$

So

$$|\hat{d}(x, y) - \hat{d}(x', y')| = 0 \iff \hat{d}(x, y) = \hat{d}(x', y')$$

which shows $\tilde{d}$ is a well-defined metric on $\tilde{X}$.

• **Step 2**: Prove that $(\tilde{X}, \tilde{d})$ is complete. Let $\{[x_n]\}_n$ be a Cauchy sequence in $(\tilde{X}, \tilde{d})$ with $x_n = \{x_{n,k}\}$ being a Cauchy sequence in (X, d).
Claim: $\{[x_n]\}_n$ is convergent to some $[x]$.
To do so, it is sufficient to show that a subsequence of it is convergent. Now, picking a subsequence that is convenient, we may assume that

$$d(x_{n,k}, x_{n,\ell}) < \frac{1}{n}$$

for all $k, \ell \geq n$ Visually, we can see this sequence

$$\begin{array}{ccccccc} [x_1] & [x_2] & \dots & [x_n] & \dots & [x_m] & \dots \\ x_{1,1} & x_{2,1} & \dots & x_{n,1} & \dots & x_{m,1} & \dots \\ x_{1,2} & x_{2,2} & \dots & x_{n,2} & \dots & x_{m,2} & \dots \\ x_{1,3} & x_{2,3} & \dots & x_{n,3} & \dots & x_{m,3} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{1,n} & x_{2,n} & \dots & x_{n,n} & \dots & x_{m,n} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{1,m} & x_{2,m} & \dots & x_{n,m} & \dots & x_{m,m} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array}$$

Let $x = \{x_{n,m}\}$ be the diagonal sequence. We want to show that $\{x_{n,m}\}$ is Cauchy in X, and $\{[x_n]\} \to [x] \in (\tilde{X}, \tilde{d})$. Since

$$d(x_{n,m}, x_{n,m}) \leq d(x_{n,n}, x_{n,k}) + d(x_{n,k}, x_{m,k}) + d(x_{m,k}, x_{m,m}) \leq \frac{1}{n} + \tilde{d}([x_n], [x_m]) + \frac{1}{m} < \epsilon$$

$x = \{x_{m,m}\}$ is Cauchy in X. As a result of result, $(\tilde{X}, \tilde{d})$ is a complete metric space.

- **Step 3**: Show that (X, d) is isometric to a dense subset of $(\tilde{X}, \tilde{d})$.
  Define $i : X \to \tilde{X}$ by $x \to [\{x_n\}]$ where $x_n = x \; \forall n \in \mathbb{N}$. We see that i is an isometric embedding.
  Claim: $i(X)$ is dense in $\tilde{X}$.
  Observe, for all $[x] \in \tilde{X}$ where $x = \{x_m\}$ is a Cauchy sequence in X. Then

  $$i(x_m) \to [x] \text{ in } (\tilde{X}, \tilde{d})$$

  Steps 1 through 3 therefore demonstrate $(\tilde{X}, \tilde{d})$ is the completion of (X, d).

- **Step 4**: We now want to show this completion is unique!
  Assume $(\bar{X}, \bar{d})$ is another completion of (X, d). We want to show that $(\tilde{X}, \tilde{d})$ and $(\bar{X}, \bar{d})$ are isometric. Indeed **Insert Commutator**

  Note that the map

  $$\bar{i} \circ \tilde{i}^{-1} : \tilde{X} \to \bar{i}((X) \subset \bar{X}$$

  is distance preserving. Thus, it is Lipschitz with $Lip(\bar{i} \circ \tilde{i}^{-1}) = 1$. By homework, since $\tilde{i}(X)$ is dense in $\tilde{X}$, there exists a Lipschitz function $f : \tilde{X} \to \bar{X}$ such that

  $$f|_{\tilde{i}(X)} = \bar{i} \circ \tilde{i}^{-1} \text{ and } Lip(f) = 1$$

  Similarly, $\exists g : \bar{X} \to \tilde{X}$ such that

  $$g|_{\bar{i}(X)} = \bar{i} \circ \tilde{i}^{-1} \text{ and } Lip(g) = 1$$

  Then $f \circ g : \bar{X} \to \bar{X}$ is Lipschitz and

  $$f \circ g|_{\bar{i}(X)} = identity|_{\bar{X}}$$

  Similarly, $g \circ f|_{\tilde{i}(X)} = identity|_{\tilde{X}}$ Therefore, $g = f^{-1}$. Then for all $x, y \in \tilde{X}$, we see

  $$\bar{d}(f(x), f(y)) \leq \tilde{d}(x, y) = \tilde{d}(g(f(x)), g(f(y))) \leq \bar{d}(f(x), f(y)) \implies \bar{d}(f(x), f(y)) = \tilde{d}(x, y)$$

  Therefore, these completions possess the same topology and are unique up to isometric isomorphism.

■

1.7 Compactness

Definition 1.60. *A subset K of a metric space is called* <u>sequentially compact</u> *if every sequence in K has a convergent subsequence whose limit belongs to K.*

Example 1.61. $(\mathbb{R}, |\cdot|)$ *is not sequentially compact.* $[a, b), (a, b]$ *are also not sequentially compact.*

Theorem 1.62 (Heine-Borel). *A subset $K \subset \mathbb{R}^n$ is sequentially compact if and only if K is closed and bounded.*

Proof. Exercise ■

Proposition 1.63. *If $A \subset X$ is sequentially compact, then A is compact.*

Note 1.64. *In $(\mathbb{R}^n, |\cdot|)$, $A \subset \mathbb{R}^n$ is complete if and only if A closed. So, in $(\mathbb{R}^n, |\cdot|)$, a set A is complete and bounded if and only if A is sequentially compact. But in a general metric space (X, d), the converse is not necessarily true.*

Example 1.65. *Suppose we have a non-finite set X equipped with the discrete metric d. Then in the space (X, d), we can choose a sequence $\{x_n\}_n \subset X$ with $\{x_n\}_n$ distinct. Then it has no convergent subsequence.*

Definition 1.66. *Let I be an index set. A collection $G = \{G_\alpha : \alpha \in I\}$ of subsets of X is called a* <u>cover</u> *of a subset $A \subset X$ if*

$$A \subset \bigcup_{\alpha \in I} G_\alpha$$

Definition 1.67. *If every G_α in the cover is open, we say that $\{G_\alpha\}$ is an* <u>open cover.</u>

Definition 1.68. *A* <u>subcover</u> *of $G = \{G_\alpha : \alpha \in I\}$ is a collection $\{G_\alpha : \alpha \in I_0\}$ where $I_0 \subset I$.*

Definition 1.69. *If $G = \{G_\alpha : \alpha \in I\}$ where I is finite, then G is a* <u>finite</u> *cover.*

Definition 1.70. *For all $\epsilon > 0, A \subset X$. An* <u>$\epsilon$-net</u> *for A is a subset of X of the form*

$$\{x_\alpha : \alpha \in I\}$$

such that

$$\{B_\epsilon(x_\alpha) : \alpha \in I\}$$

Definition 1.71. *$A \subset X$ is* <u>totally bounded</u> *if it has a finite ϵ-net cover for every $\epsilon > 0$. That is, A can be covered by finite many open ϵ-balls for any $\epsilon > 0$.*

Question 1.72. *What is the relationship between totally bounded and bounded? Clearly, if a set is totally bounded, it must be bounded. The converse is not true.*

Theorem 1.73. *A set $A \subset X$ is sequentially compact if and only if A is complete and totally bounded.*

Proof. $(\rightarrow)$ We have already proven A is complete.
Assume A is not totally bounded. The $\exists \epsilon > 0$ such that A has no finite ϵ-net. We can pick $x_1 \in A$ since $\{x_1\}$ is not an ϵ-net for A. Therefore, $x_2 \in A$ such that $d(x_1, x_2) \geq \epsilon$. Since $\{x_1, x_2\}$ is not an ϵ-net for A, we can identify $x_3 \in A$ such that $d(x_i, x_3) \geq 3$ for $i = 1, 2$. Performing this iteratively, we get a sequence $\{x_n\}_{n=1}^\infty$ such that

$$d(x_n, x_m) \geq \epsilon \text{ with } n \neq m$$

So this sequence doesn't have a convergent subsequence since this sequence cannot contain a Cauchy subsequence. But this violates the hypothesis that A is sequentially compact. Contradiction!

($\leftarrow$) Suppose A is complete and totally bounded. Let $\{x_n\}$ be any sequence in A. Since A is totally bounded, there is a sequence of balls $\{B_k\}$ such that B_k has radius $\frac{1}{2^k}$ and

$$\bigcap_{i=1}^{k} B_i$$

contains infinitely many terms of $\{x_n\}$. Thus, we may pick

$$\begin{aligned}
&x_{n_1} \in B_1 \\
&x_{n_2} \in B_2 \setminus \{x_{n_1}\} \\
&\vdots \\
&x_{n_k} \in B_2 \setminus \{x_{n_1}, \dots, x_{n_{k-1}}\}
\end{aligned}$$

By doing so, we get a Cauchy subsequence $\{x_{n_k}\}_k$. Since A is complete, $\{x_{n_k}\}_k$ converges to some point in A. Therefore, A is sequentially compact. ∎

Corollary 1.73.1. *If A is sequentially compact, then A is separable.*

Proof. For all n, let A_n be a $\frac{1}{n}$-net of A. Let $A_\infty = \bigcap_n A_n$, then A_∞ is both countable and dense in A. ∎

Definition 1.74. *A set K is compact if every open cover K has a finite subcover.*

Note 1.75. *This definition of compactness does not require K be a subset of a metric space.*

Theorem 1.76. *For metric spaces, sequential compactness is equivalent to (topological) compactness.*

Proof. ($\rightarrow$) We want to prove compactness implies sequential compactness.
Let K be a compact subset of a metric space X. Assume K is not sequentially compact. Then there is a sequence $\{y_n\} \subset K$ which has no convergent subsequence. Then for every $x \in K$, x is not a limit point of any subsequence of $\{y_n\}$. So there exists an $\epsilon_x > 0$ such that

$$B_{\epsilon_x}(x) \cap \{y_n : n \in \mathbb{N}\} = \emptyset \text{ or } \{x\}$$

Now the set

$$\{B_{\epsilon_x}(x) : x \in K\}$$

is an open cover of the compact set K. Since K is compact, we can identify a finite subcover

$$K \subset \bigcup_{i=1}^{n} B_{\epsilon_{x_i}}(x_i)$$

In particular, $\{y_n\} \subset \bigcup_{i=1}^{n} B_{\epsilon_{x_i}}(x_i)$. Thus, at least one of the balls contains infinitely many element-terms of $\{y_n\}$. But then $\{y_n\}$ has a convergent subsequence to a center x_i for some i. Contradiction! Therefore, K is sequentially compact.

∎

We pause the proof to discuss a useful concept for the next proof.

Definition 1.77. *Let $\{G_\alpha : \alpha \in I\}$ be an open cover of a subset K of a metric space X. A real number $\lambda > 0$ is called a Lebesgue number of the cover if for every ball of radius λ in K is contained in one of the sets G_α.*

Note 1.78. *Every subset of K of diameter 2λ would serve to replace the condition of every ball of radius λ in the Lebesgue number definition.*

Example 1.79. *(X, d) is a discrete metric space. Then $\lambda = \frac{1}{2}$ is a Lebesgue number for any open cover of X. (Not a unique choice)*

Example 1.80. *$X = (0, \infty)$. Define the open cover*

$$G_n := (n - \frac{1}{n}, n + 1 + \frac{1}{n})$$

for $n \in \mathbb{N}$. But the set

$$[n + 1 - \frac{1}{n+1}, n + 1 + \frac{1}{n+1}]$$

cannot be contained in a single G_n. Therefore, no such number can serve as a Lebesgue number for this cover.

Proposition 1.81. *Let K be a sequentially compact subset of a metric space X. Then for any open cover $\{G_\alpha : \alpha \in I\}$ of K has a Lebesgue number.*

Proof. Leaving this as an exercise. ■

Proof. (Returning to the proof of the previous theorem).
($\leftarrow$) Let $\{G_\alpha : \alpha \in I\}$ be any open cover of K. Since K is sequentially compact, the open cover has a Lebesgue number $\lambda > 0$. Also, since K is totally bounded, it has a finite λ-net:

$$\{x_1, \ldots, x_n\}$$

That is,

$$K \subset \bigcup_{i=1}^{n} B_\lambda(x_i)$$

But

$$B_\lambda(x_i) \cap K \subset G_{\alpha_i}$$

for some $\alpha_i \in I$. That is,

$$K \subset \bigcup_{i=1}^{n} G_{\alpha_i}$$

So $\{G_\alpha : \alpha \in I\}$ has a finite subcover $\{G_{\alpha_i}\}_i$. Therefore, K is compact! ■

In summary:

compact $\iff$ sequentially compact $\iff$ totally bounded and complete $\iff$ closed and precompact

Definition 1.82. *A set $A \subset X$ is precompact if $\overline{A}$ is compact.*

Proposition 1.83. *A is compact if and only if A is closed and precompact.*

Proof. ($\rightarrow$) A compact implies A complete implies A closed implies $\bar{A} = A$ implies A compact.

■

2 Sequences of Continuous Functions

2.1 Convergence of Functions

Let X be a metric space, and

$$f_n : X \to \mathbb{R}$$

is a function on X for all n.

- **Pointwise Convergence:** denoted $f_m \to f$ on X if

  $$f_n(x) \to f(x)$$

  for all $x \in X$. However, this has a major drawback, in that f_n continuous doesn't implies f being continuous.

  Example 2.1. $f_(x) = x^n$ *on* $[0, 1]$

- **Metric Convergence:** Introducing a suitable distance d between functions

  $$f_n \to f \iff d(f_n, f) \to 0$$

  Typically, $d(f_n, f) = \|f_n - f\|$ for some norm $\|\cdot\|$.

Example 2.2. *How to describe a function g is close to zero?*

1. *We can say the maximum is small:*

   $$\|g\|_\infty = \sup_{x \in X} |g(x)|$$

2. *We can say the "area" is small*

   $$\|g\|_1 = \int_x |g(x)| \ dx$$

   $$\|g\|_p = \left(\int_x |g(x)|^p \ dx \right)^{\frac{1}{p}}$$

3. *Sometimes, we may want control of the derivatives:*

   $$\|g\|_{W_{1,|\alpha|}} = \|g\|_p + \|D^\alpha g\|_p$$

   which serves as the norm in Sobolev Spaces.

Now, we only focus on the uniform norm / sup norm, where

$$\|f\|_\infty = \sup_{x \in X} |f(x)|$$

Definition 2.3. *A sequence $f_n : X \to \mathbb{R}$ is uniformly continuous to $f : X \to \mathbb{R}$ if*

$$\|f_n - f\|_\infty \to 0$$

often denoted $f_n \rightrightarrows f$.

Example 2.4. *The function sequence $f_n(x) = x^n$ defined on $[0, 1]$ does not converge uniformly to*

$$f(x) = \begin{cases} 1 & x \in [0, 1) \\ 0 & x = 1 \end{cases}$$

Clearly,

$$\|f_n - f\|_\infty = 1 \not\to 0$$

Remark 2.5. *If $f_n \to f$ **pointwise**, then $\exists \epsilon > 0$ such that for all $x \in X$, there exists an $N_{\epsilon,x}$ such that*

$$|f_n(x) - f(x)| \leq \epsilon$$

whenever $n \geq N$.
*On the other hand, if $f_n \to f$ **uniformly**, the for all $\epsilon > 0$, there exists an N_ϵ such that*

$$|f_n(x) - f(x)| \leq \epsilon$$

whenever $n \geq N$, regardless of $x \in X$!

2.2 The Spaces of Bounded and Continuous Functions

Definition 2.6. *Let X be a metric space, and define*

$$B(X) := \{f : X \to \mathbb{R} : f \text{ is bounded } \}$$

Then $B(X)$ is a vector space.

Proposition 2.7. *$\|\cdot\|_\infty$ is a norm on $B(X)$*

Proof.
- Proving $\|f\|_\infty = 0 \iff f = 0$ and $\|\lambda f\|_\infty = |\lambda| \, \|f\|_\infty$
- For the triangle inequality, we observe

$$\begin{aligned} \|f + g\|_\infty &= \sup_x |f(x) + g(x)| \\ &\leq \sup_x |f(x)| + \sup_x |g(x)| \\ &= \|f\|_\infty + \|g\|_\infty \end{aligned}$$

∎

Definition 2.8. *A complete, normed linear space is called a Banach Space.*

Theorem 2.9. *$(B(X), \|\cdot\|_\infty)$ is a Banach Space.*

Proof. Let $\{f_n\}_n$ be a Cauchy sequence in $B(X)$.

- We need to show that there is a function this converges to. Clearly, the only candidate is an f such that

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

  But we need to show this limit exists.
  Thus for all $\epsilon > 0$, there exists $N > 0$ such that for all $m, n \geq N$

  $$\|f_n - f_m\|_\infty < \epsilon$$

  $$\iff |f_n(x) - f_m(x)| \leq \epsilon$$

  Therefore, $\{f_n(x)\}_n$ us a Cauchy sequence in $\mathbb{R}$, which is complete. Therefore, $\lim_{n\to\infty} f_n(x)$ exists.

- Next we need to show convergence to this function f. Since we already said for all x,

  $$|f_n(x) - f(x)| \leq \epsilon$$

  for all $n \geq N$, so $\|f - f_n\|_\infty \leq \epsilon \implies f_n \rightrightarrows f$.

- Lastly, we must show that $f \in B(X)$. So for $\epsilon = 1$, $n = N(\epsilon)$, then

$$|f(x)| \leq |f_N(x) - f(x)| + |f_N(X)| \leq 1 + |f_N(x)|$$

since $f_N \in B(X)$, then f must also be bounded.

■

Question 2.10. *Is $(B(X), \|\cdot\|_\infty$ compact? No! Consider the sequence defined*

$$f_n(x) = n$$

Then $\{f_n\}_n$ has no convergent subsequence.

Consider a new set

$$C(X) := \{f : X \to \mathbb{R} : f \text{ is continuous }\}$$

Note 2.11. *$\|\cdot\|_\infty$ is not necessarily a norm on $C(X)$ since $\|f\|_\infty \to \infty$.*

So we let

$$C_b(X) = C(X) \cap B(X) := \{f : X \to \mathbb{R} : f \text{ is continuous and bounded }\}$$

Proposition 2.12. *$C_b(X)$ is a closed subset of $(B(X), \|\cdot\|_\infty$*

Proof. We want to show that if $f_n \to f$ uniformly with $f_n \in C_b(X)$ then $f \in C_b(X)$. **Complete Proof** ■

Since $B(X)$ is a Banach space, it follows that $(C_b(X), \|\cdot\|)$ is a Banach space for any metric space X.

Question 2.13. *If X is compact, then what is the relationship between $C(X)$ and $B(X)$? Recall from the text that f continuous defined on a compact space must achieve both its maximum and minimum, and therefore*

$$C(X) \subset B(X) \implies C_b(X) = C(X)$$

Corollary 2.13.1. *If X is compact, then $(C(X), \|\cdot\|_\infty)$ is Banach.*

Example 2.14. *Suppose $X = \{x_1, \ldots, x_n\}$ is a finite set equipped with the discrete metric d. A function $f : X \to \mathbb{R}$ can be identified with a point $\{y_1, y_2, \ldots, y_n\}$ with $y_i = f(x_i)$. Every function on (X, d) is contiuous. Therefore*

$$(C(X), \|\cdot\|_\infty) \cong (\mathbb{R}^n, \|\cdot\|_\infty)$$

Therefore, by the previous corollary, $(\mathbb{R}^n, \|\cdot\|_\infty)$ is complete.

Definition 2.15. *The support of a function $f : X \to \mathbb{R}$ is given by*

$$supp(f) := \overline{\{x \in X : f(x) \neq 0\}}$$

Example 2.16. *If*

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

Then $supp(f) = \mathbb{R}$

Note 2.17. *$supp(f)$ is closed.*

Note 2.18. *$f|_{X \setminus supp(f)} = 0$.*

We define the set

$$C_c(X) := \{f \in C(X) : supp(f) \text{ is compact}\} \subset C(X)$$

Question 2.19. *Is $C_c(X)$ a closed subset of $C_b(X)$? No! You can create a sequence of bump functions $f_n \in C_c(X)$ such that $f_n \rightrightarrows f \notin C_c(X)$.*

Now define:

$$C_0(X) := \overline{C_c(X)}$$

and equip this set with the $\|\cdot\|_\infty$. We see this set fits within the chain:

$$C_c(X) \subset C_0(X) \subset C_b(X) \subset C(X) \cap B(X)$$

Note 2.20. *$(C_0(X), \|\cdot\|_\infty)$ is a Banach space since it's a closed subset of a complete space!*

Proposition 2.21. *$f \in C_0(X)$ if and only if for every $\epsilon > 0$, there exists $K \subset X$ compact such that*

$$x \notin K \implies |f(x)| < \epsilon$$

Proof. For f in $C_0(X)$, we can find $f_x \in C_c(X)$ such that $f_n \rightrightarrows f$. That is for all $\epsilon > 0$ there exists N such that

$$\|f_n - f\|_\infty < \epsilon$$

for all $n \geq N$. Now take $K = supp(f_n)$. Then for all $x \in K, f_N(x) = 0$. So

$$|f(x)| = |f(x) - f_N(x)| < \epsilon$$

■

2.3 Polynomial Approximations

Definition 2.22. *The Bernstein Polynomials. We define a basis of unity:*

$$\mathcal{X}_{n,k}(x) = \binom{n}{k} x^k (1-x)^{n-k}$$

Then by the binomial theorem:

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

So for any $f \in C([0,1])$, the Bernstein Polynomial of f is

$$B_n(x; f) := \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \mathcal{X}_{n,k}(x)$$

Example 2.23. *For any constant function:*

$$B_n(x; c) = c = c \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \mathcal{X}_{n,k}(x)$$

Example 2.24.

$$B_n(x; x) = \sum_{k=0}^{n} \left(\frac{k}{n}\right) \mathcal{X}_{n,k}(x)$$

Differentiating a basis function with respect to x:

$$\begin{aligned} n &= \sum_{k=0}^{n} \binom{n}{k} k x^{k-1}(1-x)^{n-k} \\ \iff x &= \sum_{k=0}^{n} \binom{n}{k} \frac{k}{n} x^k (1-x)^{n-k} \\ &= \sum_{k=0}^{n} \frac{k}{n} \mathcal{X}_{n,k}(x) \\ &= B_n(x;x) \end{aligned}$$

As a result, for any linear function f,

$$B_n(x,f) = f$$

Taking the next derivative with respect to x:

$$\begin{aligned} n(n-1) &= \sum_{k=0}^{n} \binom{n}{k} k(k-1) x^{k-2}(1-x)^{n-k} \\ \iff \frac{n(n-1)x^2}{n^2} &= \sum_{k=0}^{n} \binom{n}{k} \frac{k(k-1)}{n^2} x^k (1-x)^{n-k} \\ &= \sum_{k=0}^{n} \binom{n}{k} \left(\frac{k^2}{n^2} - \frac{k}{n^2}\right) x^k (1-x)^{n-k} \\ &= \sum_{k=0}^{n} \binom{n}{k} \left(\frac{k^2}{n^2}\right) x^k (1-x)^{n-k} - \sum_{k=0}^{n} \binom{n}{k} \left(\frac{k}{n^2}\right) x^k (1-x)^{n-k} \\ &= \sum_{k=0}^{n} \left(\frac{k^2}{n^2}\right) \mathcal{X}_{n,k}(x) - \sum_{k=0}^{n} \left(\frac{k}{n^2}\right) \mathcal{X}_{n,k}(x) \\ &= B_n(x;x^2) - \frac{1}{n} B_x(x;x) \end{aligned}$$

Therefore,

$$B_x(x;x^2) = (1-\frac{1}{n})x^2 + \frac{1}{n}x = x^2 + \frac{1}{n}(x-x^2)$$

Theorem 2.25 (Weierstrass Approximation Theorem)**.** *The set of polynomials $P([0,b])$ is dense in $(C(X), \|\cdot\|_\infty)$*

Proof. Without loss of generality, we can prove on the set $[0,1]$. We want to show that

$$\|B_x(x;f) - f\|_\infty \to 0$$

Observe:

$$\begin{aligned} |B_x(x;f) - f(x)| &= \left| \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \mathcal{X}_{n,k}(x) - f(x) \right| \\ &= \left| \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \mathcal{X}_{n,k}(x) - f(x) \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \right| \\ &\le \sum_{k=0}^{n} \left| f\left(\frac{k}{n}\right) - f(x) \right| \mathcal{X}_{n,k}(x) \end{aligned}$$

Since f is continuous on the compact set $[0,1]$, it is uniformly continuous. Thus for every $\epsilon > 0$, there is a $\delta > 0$ such that

$$|f(x) - f(y)| < \epsilon$$

whenever $x, y \in [0,1], |x - y| < \delta$. Fixing $x \in [0,1]$, with $n \geq N$, then we let

$$I_x := \{k : |\frac{k}{n} - x| < \delta\}$$

$$J_x := \{k : |\frac{k}{n} - x| \geq \delta\}$$

Therefore, we can split our summation as follows:

$$\begin{aligned}
|B_x(x; f) - f(x)| &\leq \sum_{k=0}^{n} \left| f\left(\frac{k}{n}\right) - f(x) \right| \mathcal{X}_{n,k}(x) \\
&= \sum_{k \in I_x} \left| f\left(\frac{k}{n}\right) - f(x) \right| \mathcal{X}_{n,k}(x) + \sum_{k \in J_x} \left| f\left(\frac{k}{n}\right) - f(x) \right| \mathcal{X}_{n,k}(x) \\
&\leq \epsilon \sum_{k \in I_x} \mathcal{X}_{n,k}(x) + 2 \|f\|_\infty \sum_{k \in J_x} \mathcal{X}_{n,k}(x) \\
&\leq \epsilon + 2 \|f\|_\infty \sum_{k \in J_x} \frac{(x - \frac{k}{n})^2}{\delta^2} \mathcal{X}_{n,k}(x) \\
&\leq \epsilon + \frac{2 \|f\|_\infty}{\delta^2} \sum_{k \in J_x} (x - \frac{k}{n})^2 \mathcal{X}_{n,k}(x) \\
&= \epsilon + \frac{2 \|f\|_\infty}{\delta^2} \sum_{k \in J_x} \left(x^2 - \frac{2kx}{n} + (\frac{k}{n})^2 \right) \mathcal{X}_{n,k}(x) \\
&= \epsilon + \frac{2 \|f\|_\infty}{\delta^2} \left(x^2 - 2x B_n(x; x) + B_n(x; x^2) \right) \mathcal{X}_{n,k}(x) \\
&= \epsilon + \frac{2 \|f\|_\infty}{\delta^2} \left(x^2 - 2x \cdot x + x^2 + \frac{1}{n}(x - x^2) \right) \mathcal{X}_{n,k}(x) \\
&= \epsilon + \frac{2 \|f\|_\infty}{\delta^2} \cdot \frac{1}{n}(x - x^2) \\
&\leq \epsilon + \frac{\|f\|_\infty}{4n\delta^2} \leq 2\epsilon
\end{aligned}$$

Therefore, with a choice of n large enough, $B_n(x; f) \to f$ uniformly! ■

2.4 Compactness Implications

Definition 2.26. *A metric space X space is Hausdorff if for every $x \neq y$, then we can identify open subsets $U_x, V_y \subset X$ such that*

$$x \in U_x, y \in V_y \text{ and } U_x \cap V_y = \emptyset$$

We then say A can separate points if we can identify an $a \in A$ "between" two points.

Note 2.27. *Recall that a set $A \in \mathbb{R}^n$ is compact if and only if A is closed and bounded. What can be said about the set a compact set $\mathcal{F}$ such that*

$$\mathcal{F} \subset (C(K), \|\cdot\|_\infty)?$$

Then $\mathcal{F}$ is closed and bounded. Is the converse true?

Example 2.28. *We define a sequence of functions by continuous pyramids,*

$$f_n(x) = \begin{cases} 0 & x < \frac{1}{2^{n+1}} \\ 1 & x = \frac{\frac{1}{2^{n+1}} + \frac{1}{2^n}}{2} \\ 0 & x > \frac{1}{2^n} \end{cases}$$

Clearly f_n is bounded since

$$\|f - 0\|_\infty = 1$$

However, if $n \neq m$, then

$$\|f_n - f_m\|_\infty = 1$$

so $\{f_n\}_n$ has no convergent subsequence. Therefore, not compact. But this set is closed. That is, this is closed and bounded but not compact!!!

Note 2.29. *$\mathcal{F} \in C(K)$ is bounded means that for all $f \in$, there exists both $g \in C(X)$ and $M \geq 0$ such that*

$$\|f - g\|_\infty \leq M$$

So

$$\|f\|_\infty \leq \|f - g\|_\infty + \|g\|_\infty \leq M + \|g\|_\infty$$

We use this to motivate a definition.

Definition 2.30. *Let $\mathcal{F}$ be a family of functions*

$$f : (X, d_X) \to (Y, d_Y)$$

between metric spaces. Then $\mathcal{F}$ is called equibounded provided for all $f \in \mathcal{F}$ there exists $g \in \mathcal{F}$ if

$$\|f\|_\infty \leq M + \|g\|_\infty$$

Corollary 2.30.1. *$\mathcal{F}$ is bounded means every function in $\mathcal{F}$ is equibounded.*

Definition 2.31. *Let $\mathcal{F}$ be a family of functions*

$$f : (X, d_X) \to (Y, d_Y)$$

between metric spaces. Then $\mathcal{F}$ is called equicontinuous at $x \in X$ if for every $\epsilon > 0$, there is a $\delta = \delta(\epsilon, x) > 0$ such that

$$d_X(x, y) < \delta \implies d_Y(f(x), f(y)) \leq \epsilon$$

for all $f \in \mathcal{F}$.

Note 2.32. *δ is independent of the function f. If δ is also independent of $x \in X$, then $\mathcal{F}$ is uniformly equi-continuous.*

Example 2.33.

$$Lip_M(X) := \{f : X \to \mathbb{R} : Lip(f) \leq M\}$$

is equi-continuous with $\delta = \frac{\epsilon}{M}$.

Proposition 2.34. *Let X be a compact metric space, and*

$$\mathcal{F} \subset C(X, Y) := \{f : X \to Y : f \text{ continuous}\}$$

Then $\mathcal{F}$ is equi-continuous. Moreover, $\mathcal{F}$ is uniformly equi-continuous.

Proof. Assume then there is some $\epsilon > 0$ such that for each n, there exists $x_n, y_n \in X$ with

$$d_X(x_n, y_n) < \frac{1}{n} \quad \text{but} \quad d_Y(f_n(x_n), f_n(y_n)) \geq \epsilon$$

for some $f_n \in \mathcal{F}$. Since X is compact, by taking a subsequence if necessary, we may assume that

$$x_n \to x^* \quad y_n \to y^*$$

in X for some $x^*, y^* \in X$. Since

$$d_X(x^*, y^*) = \lim_n d_X(x_n, y_n) = 0$$

we have $x^* = y^*$. For x^*, since $\mathcal{F}$ is equi-continuous, then there is δ such that

$$d_X(x, x^*) < \delta \implies d_Y(f(x), f(x^*)) < \frac{\epsilon}{2}$$

for all $f \in \mathcal{F}$. Now when n is large enough,

$$d_X(x_n, x^*) < \delta, d_X(y_n, y^*) < \delta$$

Thus,

$$d_Y(f(x_n), f(y_n)) \leq d_Y(f(x_n), f(x^*)) + d_Y(f(y_n), f(y^*)) < \epsilon$$

This contradicts our assumption! ■

Theorem 2.35 (Stone-Weierstrass Theorem). *Suppose X is a compact Hausdorff space, and A is a subalgebra of $C(X, \mathbb{R})$ which contains a non-zero constant function. Then A is dense in $C(X, \mathbb{R})$ if any only if it separates points.*

Proof. **Do on your own time?** ■

Proposition 2.36. *Let K be a compact metric space. If a sequence $\{f_n\}_{n=1}^\infty \subset C(K)$ is bounded and equi-continuous, then it has a (uniformly) convergent subsequence.*

Proof. Sketch of Proof:

1. K has a countable dense subset.
2. Find a subsequence $\{g_n\}_{n=1}^\infty$ of $\{f_n\}_{n=1}^\infty$ that converges pointwise on S via diagonal argument
3. Show that $\{g_n\}_{n=1}^\infty$ is a Cauchy sequence in $C(K)$.

Actual proof:

1. Since K is compact, it is seperable. Thus, it has a countable dense subset $S = \{x_1, x_2, \ldots\}$.
2. Since $\{f_n\}_{n=1}^\infty$ is bounded in $C(K)$, then sequence $\{f_n(x_1)\}_{n=1}^\infty$ is bounded in $\mathbb{R}$. Thus, we may choose a subsequence $\{f_{1,n}\}_{n=1}^\infty \subset \{f_n\}_{n=1}^\infty$ such that

   $$\{f_{1,n}(x_1)\}_{n=1}^\infty \text{ converges}$$

   Similarly, we may choose a subsequence $\{f_{2,n}\}_{n=1}^\infty$ such that $\{f_{2,n}(x_2)\}_{n=1}^\infty$ converges. Repeating this process, we obtain sequences

   $$\{\{f_{k,n}\}_{n=1}^\infty\}_{k=1}^\infty$$

   such that

   $$\{f_{k,n}(x_k)\}_{n=1}^\infty \text{ converges}$$

   for every $1 \leq k$. Finally, we define a diagonal subsequence g_k by

   $$g_k = f_{k,k}$$

   Notice,

   $$\{g_{k,k}(x_i)\}_{k=1}^\infty \text{ converges}$$

   for each $x_i \in S$!

3. Since $\{g_{k,k}(x_i)\}_{k=1}^{\infty}$ is also equi-continuous in K, $\{g_{k,k}(x_i)\}_{k=1}^{\infty}$ is uniformly equi-continuous. Thus, for all $\epsilon > 0$, there exists $\delta > 0$ such that

$$d(x,y) < \delta \implies |g_k(x) - g_k(y)| < \frac{\epsilon}{3}$$

for all k. Since S is dense in K, $\{B_\delta(x) : x \in S\}$ is any open cover of K. K is compact implies there must be a finite subcover in S, such that

$$K \subset \bigcup_{i=1}^{\ell} B_\delta(x_i)$$

since $\{g_{k,k}(x_i)\}_{k=1}^{\infty}$ is convergent for each $1 \le i \le \ell < \infty$, then there must exists $N > 0$ such that for all $m, n \ge N$, then

$$g_m(x_i) - g_n(x_i)| < \frac{\epsilon}{3}$$

Thus, for all $x \in K \subset \bigcup_{i=1}^{\ell} B_\delta(x_i)$, then there must exists i such the

$$x \in B_\delta(x_i)$$

Therefore, when $m, n \ge N$, then

$$\begin{aligned} |g_m(x) - g_n(x)| &\le |g_m(x) - g_m(x_i)| + |g_m(x_i) - g_n(x_i)| + |g_n(x_i) - g_n(x)| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon \end{aligned}$$

Theefore,

$$\|g_m - g_n\| < \epsilon$$

and $\{g_{k,k}(x_i)\}_{k=1}^{\infty}$ is Cauchy in $C(K)$.Moreover, since $(C(K), \|\cdot\|)$ is complete, then $\{g_{k,k}(x_i)\}_{k=1}^{\infty}$ is uniformly convergent.

■

Recall $A \subset X$ is precompact if $\bar{A}$ is compact. Then

$$A \text{ is compact} \iff A \text{ is closed and precompact}$$

Proposition 2.37. *Let K be a compact metric space. Then*

$$\mathcal{F} \subset C(K) \text{ precompact} \iff \mathcal{F} \text{ is bounded and equi-continuous}$$

Proof. ($\Leftarrow$) Follows from the previous proposition.
($\Rightarrow$) For every $\epsilon > 0$, since $\overline{\mathcal{F}}$ is compact, so it has a finite $\frac{\epsilon}{6}$-net. Since $\mathcal{F} \subset \overline{\mathcal{F}}$, it also has a finite $\frac{\epsilon}{6}$-net of $\mathcal{F}$. By the homework problem, there is an $\frac{\epsilon}{3}$-net of $\mathcal{F}$ whose centers are in $\mathcal{F}$. That is

$$\mathcal{F} \subset \bigcup_{i=1}^{\infty} B_{\frac{\epsilon}{3}}(f_i)$$

for some $f_i \in \mathcal{F}, 1 \le i \le n$. Clearly, $\mathcal{F}$ must be bounded.

Since K is compact, each continuous function f_i is uniformly continuous on K. Thus, there exists $\delta_i > 0$ such that

$$|f_i(x) - f_i(y)| < \frac{\epsilon}{3}$$

whenever $d(x,y) < \delta_i$. Let

$$\delta := \min_{1 \le i \le n} \delta_i > 0$$

Then for every $f \in \mathcal{F}$, $f \in B_{\frac{\epsilon}{3}}(f_i)$ for some i. That is

$$\|f - f_i\|_\infty < \frac{\epsilon}{3}$$

Then whenever $d(x, y) \leq \delta \leq \delta_i$, we have

$$|f(x) - f(y)| \leq |f_i(x) - f(x)| + |f_i(x) - f_i(y)| + |f_i(y) - f(y)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

Therefore, $\mathcal{F}$ is equi-continuous. ■

Theorem 2.38 (Arzela-Ascoli)**.** *Let K be a compact metric space. A subset $\mathcal{F} \subset C(K)$ is compact if and only if it is closed, bounded, and equicontinuous.*

Example 2.39. *Let $K := \{x_1, x_2, \ldots, x_n\}$ equipped with the discrete metric. Then*

$$(C(K), \|\cdot\|_\infty) \cong (\mathbb{R}^n, \|\cdot\|_\infty)$$

A subset $\mathcal{F}$ is equicontinuous means for every $\epsilon, \exists \delta$ such that for all $f \in \mathcal{F}$

$$d(x_i, x_j) < \delta \implies |f(x_i) - f(x_j)| < \epsilon$$

Take $\delta = \frac{1}{2}$, then the above is always true. That is, for any subset of $C(K)$ is automatically equicontinuous. Thus, in $\mathbb{R}^n$, $A \subset \mathbb{R}^n$ is compact if and only if it's closed and bounded, which serves as a quick proof the Heine-Borel theorem.

Example 2.40. *Let K be compact. Then*

$$Lip_M(K) \text{ is equicontinuous with } \delta = \frac{\epsilon}{M}$$

Therefore,

$$f_n \rightrightarrows f \implies f \in Lip_M(K) \implies Lip_M(K) \text{ is closed}$$

However, $Lip_M(K)$ is not bounded. By the characterization of Arzela-Ascoli, then $Lip_M(K)$ is not compact. However, any bounded and closed subset of $Lip_M(K)$ is compact.

3 Topological Spaces

3.1 Motivation and Definition

On $\mathbb{R}^n$, we have defined many matrics (e.g. $\|\cdot\|_p, 1 \le p \le \infty$). However $x_n \to x$ is the same in all these metrics. Moreover, continuity and compactness are also the same between these spaces. But we need to abstract upon their language in order to show this equivalence.

Question 3.1. *How to study 'convergence', 'sequentially compact' in a space more general than metric spaces?*

Note 3.2. *In metric spaces,*

$$\textit{sequentially compact} \iff \textit{compact i.e. open subcover exists}$$

$$\textit{continuity} \iff f^{-1}(U) \textit{ open whenever } U \textit{ open}$$

Idea!: We can replace the distance function by a collection of open sets. So, how do we define "open sets" in X?

Definition 3.3. *Let X be a non-empty set. A collection $\mathcal{T}$ os subsets of X is a topology on X if*

1. $\emptyset, X \in \mathcal{T}$
2. Closed under arbitrary unions*: If $G_\alpha \in \mathcal{T}$ for any $\alpha \in I$, then*
$$\bigcup_\alpha G_\alpha \in \mathcal{T}$$
3. Closed under finite intersections*: If $G_i \in \mathcal{T}$ for any $1 \le i \le n$, then*
$$\bigcap_{i=1}^n G_i \in \mathcal{T}$$

We call $(X, \mathcal{T})$ a topological space, and each element of $G \in \mathcal{T}$ is called an open set.

Definition 3.4. *$A \subset X$ is closed if A^c is open.*

Example 3.5. *Let X be a metric space. Then the collection of all "open sets" in X form a topology on X, called the metric topology.*

Definition 3.6. *A topological space $(X, \mathcal{T})$ is called metrizable if $\mathcal{T}$ is the metric topology derived from a metric on X.*

Example 3.7. *Let X be any nonempty set. and let*

$$2^X := \{ \textit{ subsets of } X\}$$

- *Trivial Topology* $\{\emptyset, X\}$
- *Discrete Topology* $\mathcal{T} = 2^X$ *Notice that all subsets of X in this topology are both open and closed.*
- *Let $A \subset X$, and $\mathcal{T}$ is a topology on X, then*
$$\mathcal{T}|_A := \{G \cap A : G \in \mathcal{T}\}$$
is a tpology of A, called the induced or relative topology.

Example 3.8. $(0, 1]$ *is relatively open in* $[-1, 1]$.

- $X = \{a, b, c\}$. *How many topologies can be identified on this set?*

$$\mathcal{T} = \{\{a\}, \{b, c\}, \{a, b, c\}, \emptyset\}$$

- *If* X *is a nonempty set, then we can have*

$$\mathcal{T} = \{X \setminus C : C \textit{is finite subset of } X\} \cup \{\emptyset\}$$

Definition 3.9. *A set* $V \subset X$ *is a neighborhood of* $x \in X$*, if there exists an open set* $G \subset V$ *with* $x \in G$.

Definition 3.10. *A topology* $\mathcal{T}$ *is called Hausdorff if for any* $x, y \in X$*,* $x \neq y$*, there exists a neighborhood* V_x *of* x *and* V_y *of* y *such that*

$$V_x \cap V_y = \emptyset$$

Example 3.11. *All metric topologies are Hausdorff. The discrete topology is Hausdorff, but the trivial topology is not.*

3.2 Convergence and Continuity

Definition 3.12. *A sequence* $\{x_n\}_n$ *in* X *converges to a limit* $x \in X$ *if for every neighborhood* V_x *of* x*, there exists* N *such that that*

$$x_n \in V_x$$

for all $n \geq N$.

Definition 3.13. *A function* $f : X \to Y$ *is continuous if for every neighborhood* W *of* $f(x)$*, there must exists a neighborhood* V *of* x *such that*

$$f(V) \subset W$$

Theorem 3.14. *Let* $(X, \mathcal{T})$ *and* $(Y, \mathcal{S})$ *be two topological spaces, and* $f : X \to Y$*. Then the following statements are equivalent:*

- f *is continuous on* X.
- *for any open set* $G \subset Y$*, then* $f^{-1}(G)$ *is open in* X.
- *for any closed set* $F \subset Y$*, then* $f^{-1}(F)$ *is closed in* X.
- *for any* $A \subset X$*, then* $f(\overline{A}) \subset \overline{f(A)}$
- *for any* $B \subset Y$*, then* $\overline{f^{-1}(B)} \subset f^{-1}(\overline{B})$

Here, the closure of $\overline{A}$

$$\overline{A} := \bigcap_{A \cap U^c \neq 0} U^c$$

Definition 3.15. *A function* $f : X \to Y$ *between topological spaces is called a homeomorphism if* f *is bijective and both* f *and* f^{-1} *are continuous.* X *and* Y *are called homeomorphic if there exists a homeomorphism* $f : X \to Y$

Note 3.16. *If* f *is homeomorphism, then*

$$G \subset X \textit{ open} \iff f(G) \subset Y \textit{ is open}$$

$$x_n \to x \iff f(x_n) \to f(x)$$

So homeomorphic spaces are indistinguishable as topological spaces.

Example 3.17. $([0, 1], |\cdot|)$ *and* $([0, 5], |\cdot|)$ *are homeomorphic.*

Example 3.18. *Unit line* $[0, 2\pi)$ *and unit circle* $\mathbb{S}^1$ *are not homeomorphic since*

$$\theta \to (\cos(\theta), \sin(\theta))$$

is not a homeomorphism

3.3 Base of a Topology

Definition 3.19. *Let $(X, \mathcal{T})$ be a topological space. $\mathcal{B} \subset \mathcal{T}$ is a base for $\mathcal{T}$ if for all $G \in \mathcal{T}$, there exists a collection of sets $\alpha, B_\alpha \in \mathcal{B}$ such that*

$$G = \bigcup_\alpha B_\alpha$$

Therefore, the set $\mathcal{T}$ is determined by its subset $\mathcal{B}$, where

$$\mathcal{T} = \{ \bigcup_{B \in \tilde{B}} B : \tilde{B} \subset \mathcal{B} \}$$

Example 3.20. *If $X = (\mathbb{R}, |\cdot|)$, then*

$$\mathcal{B} = \{(a, b) : a < b\}$$

Example 3.21. *If $\mathcal{T}$ is the discrete topology then on X, then*

$$\mathcal{B} = \{\{x\} : x \in X\}$$

Proposition 3.22. *Let $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ be two topological spaces, and $\mathcal{B}$ is a base for $\mathcal{S}$. Then $f : X \to Y$ is continuous if and only if $f^{-1}(B) \in \mathcal{T}$ for all $B \in \mathcal{B}$.*

Proposition 3.23. *Let $\mathcal{B}$ be a base for $\mathcal{T}$. Then $A \subset X$ compact if and only if every open cover consisting of members of $\mathcal{B}$ has a finite subcover.*

Question 3.24. *Given $\mathcal{B} \subset \mathcal{T}$, how to check if $\mathcal{B}$ is a base of $\mathcal{T}$?*

Theorem 3.25. *Suppose $\mathcal{B} \subset \mathcal{T}$ Then $\mathcal{B}$ is a base of $\mathcal{T}$ if and only if for all $x \in X$, and any neighborhood U_x of x, there exists $V_x \in \mathcal{B}$ such that*

$$x \in V_x \subset U_x$$

Proof. ($\leftarrow$) For any $U \in \mathcal{T}$, and any $x \in U$, there exists $V_x \in \mathcal{B}$ such that $x \in V_x \subset U_x$. Then

$$U = \bigcup_{x \in U} V_x$$

($\to$) Suppose $\mathcal{B}$ is a base of $\mathcal{T}$. Then for any $x \in X$ and any neighborhood U_x of x. Then there exists an open set $W_x \in \mathcal{T}$ such that $x \in W_x \subset U_x$. Since $\mathcal{B}$ is a base,

$$x \in W_x = \bigcup_{B \in \tilde{B}} B$$

for some $\tilde{B} \subset \mathcal{B}$. Then there exists $B \in \tilde{B}$ such that

$$x \in B \subset W_x \subset U_x$$

∎

Definition 3.26. *For any $x \in X$, let*

$$N_x := \{ \text{ all neighborhoods of } x \in X\}$$

A subset $W_x \subset N_x$ is called a neighborhood base or a local base for any x, if for any neighborhood $V \in N_x$, there exists $W \in W_x$ such that

$$x \in V \subset V$$

Theorem 3.27. *$\mathcal{B} \subset \mathcal{T}$ is a base of $\mathcal{T}$ if and only if for any $x \in X$, $\mathcal{B}$ contains a neighborhood base for x, which is*

$$\{B \in \mathcal{B} : x \in B\}$$

Proposition 3.28. *$f : X \to Y$ is continuous at $x \in X$ if and only if $f^{-1}(V)$ is a neighborhood of x for any V from a neighborhood base of $f(x) \in Y$.*

Proposition 3.29. *$x_n \to x$ in X if and only if for any V_x in a neighborhood base of x, there exists N such that $x_n \in V_x$ whenever $n \geq N$.*

Definition 3.30. *A topological space $(X, \mathcal{T})$ is*

- *first countable if every $x \in X$ has a countable neighborhood base*
- *second countable $\mathcal{T}$ has a countable neighborhood base*

Note 3.31. *X is second countable implies X is first countable.*

Example 3.32. *$(X, \mathcal{T}_{discrete})$ is first countable always. When X is uncountable, then X is not second countable.*

Example 3.33. *Suppose X is uncountable. Consider*

$$\mathcal{T} = \{X \setminus C : C \subset X \text{ countable } \} \cup \{\emptyset\}$$

Then $(X, \mathcal{T})$ is not first countable.

Example 3.34. *Let X be a metric space, $\mathcal{T}:=$ metric topology. Then X is first countable. And when X is separable, then X is second countable.*

Definition 3.35. *Let $\phi \subset \mathcal{T}$. If the set of all finite intersections of members in ϕ*

$$\{S_1 \cap S_2 \cap \ldots \cap S_n : S_i \in \phi, 1 \leq i \leq n\}$$

forms a base of $\mathcal{T}$, then ϕ is called a subbase of $\mathcal{T}$.

Example 3.36. *$X = (\mathbb{R}, |\cdot|)$, then*

$$\phi = \{(a, \infty) : a \in \mathbb{R}\} \bigcup \{(-\infty, b) : b \in \mathbb{R}\}$$

forms a subbase.

Theorem 3.37. *Let ϕ be a collection of subsets of a nonempty set X. If $X = \bigcup_{S \in \phi} S$, then there exists a unique topology $\mathcal{T}$ on X having ϕ as its subbase. Moreover, if*

$$\mathcal{B} = \left\{ \bigcap_{i=1}^{n} S_i : S_i \in \phi, 1 \leq i \leq n \in \mathbb{N} \right\}$$

then $\mathcal{B}$ serves as a basis for $\mathcal{T}$.

3.4 Comparing Topologies

Definition 3.38 (Point-wise Topology)**.** *Let $X = \{f : [a, b] \to \mathbb{R}\}$ be the collection of all real-valued functions on $[a, b]$. The for all $x \in [a, b], y \in \mathbb{R}, \epsilon > 0$ we define*

$$B_{x,y,\epsilon} = \{f \in X : |f(x) - y| < \epsilon\}$$

Note

$$\bigcup_{x,y,\epsilon} B_{x,y,\epsilon} = X$$

So $\{B_{x,y,\epsilon}\}$ is the subbase of a topology on X. This topology is the point-wise topology $\mathcal{T}_{pointwise}$

Example 3.39. *Further, we can form a sub-base by taking finite intersections of the base elements. Given $(x_i, y_i, \epsilon_i), 1 \leq i \leq n$, then*

$$\bigcap_{i=1}^{n} B_{x_i,y_i,\epsilon_i} = \{f \in X : |f(x_i) - y_i| < \epsilon_i for\ 1 \leq i \leq n\}$$

Question 3.40. *Why use the point-wise topology?*

Suppose $\{f_n\}_n$ converges to f with respect to the point-wise topology. Then for any neighborhood V of f, there exists N such that

$$n \geq N \implies f_n \in V$$

Pick $V \in B_{x_0,f(x_0),\epsilon}$ which is a neighborhood of V. So

$$f_n \in V \implies |f_n(x) - f(x)| < \epsilon$$

for all $\epsilon > 0$. Therefore

$$f_n(x_0) \to f(x_0)$$

for all $x \in [a, b]$.

Definition 3.41 (Uniform Topology)**.** *We define the uniform topology $\mathcal{T}_{uniform}$, by defining*

$$B_\epsilon(f) := \{g : [a, b] \to \mathbb{R} : \sup_{x \in [a,b]} |g(x) - f(x)| < \epsilon\}$$

Proposition 3.42.

$$\mathcal{T}_{pointwise} \subset \mathcal{T}_{uniform}$$

Proof. For all $U \in \mathcal{T}_{\text{pointwise}}$. We want to show this neighborhood is within the uniform topology. Observe, for all $f \in U$, so there exists a basis elements such that

$$f \in \bigcup_{i=1}^{n} B_{x_i,y_i,\epsilon_i} \subset U$$

Therefore, for $1 \leq i \leq n$. Then

$$|f(x_i) - f(y_i)| < \epsilon_i$$

Therefore, choosing $0 < \epsilon \leq \min_i\{\epsilon_i - |f(x_i) - y_i|\}$. Then for all $g \in B_\epsilon(f) \in \mathcal{T}_{\text{uniform}}$

$$|g(x_i) - y_i| \leq |g(x_i) - f(x_i)| + |f(x_i) - y_i| < \epsilon + |f(x_i) - y_i| \leq \epsilon_i$$

That is, $g \in \bigcup_{i=1}^{n} B_{x_i,y_i,\epsilon_i} \subset U$ and $B_\epsilon(f) \subset U$. Therefore, $U \in \mathcal{T}_{\text{uniform}}$! ■

Example 3.43. *Let $U = \{h : [0, 2] \to \mathbb{R} : |h(x)| < g(x)\}$. Then $U \notin \mathcal{T}_{uniform}$ because*

$$f = 0 \in U, \text{ but } B_\epsilon(f) \not\subset U \forall \epsilon > 0$$

But $U = \bigcap_{x \in [0,2], \epsilon > 0} B_{x,0,g(x)} \notin \mathcal{T}_{uniform}$

We now want a way to compare topologies to one-another.

Definition 3.44. *Let $\mathcal{T}_1, \mathcal{T}_2$ be two topologies on the same space X. If $\mathcal{T}_1 \subset \mathcal{T}_2$, then we say $\mathcal{T}_2$ is finer than $\mathcal{T}_1$. On the other hand, we say $\mathcal{T}_1$ is coarser than $\mathcal{T}_2$.*

Example 3.45. *$\mathcal{T}_{uniform}$ is finer than $\mathcal{T}_{pointwise}$*

Example 3.46. *For any topology $\mathcal{T}$ on X, then*

$$\underbrace{\{\emptyset, X\}}_{\text{trivial topology (weakest)}} \subset \mathcal{T} \subset \underbrace{2^X}_{\text{discrete topology (strongest)}}$$

Proposition 3.47. *If $\mathcal{T}_1 \subset \mathcal{T}_2$, then $x_n \to x$ with respect to $\mathcal{T}_2$, then $x_n \to x$ with respect to $\mathcal{T}_1$. That is, strong convergence implies weak convergence.*

Why should this be true? Well x_n must be in more neighborhoods surrounding x in a stronger topology. This places more requirements on convergence.

Proposition 3.48. *Suppose $f : (X, \mathcal{T}_1) \to (Y, \mathcal{S}_1)$ is continuous. If $\mathcal{T}_1 \subset \mathcal{T}_2$, $\mathcal{S}_2 \subset \mathcal{S}_1$, then $f : (X, \mathcal{T}_2) \to (Y, \mathcal{S}_2)$ is also continuous.*

Example 3.49. *If $f : (X, \mathcal{T}_{discrete}) \to (Y, \mathcal{S})$ then f is always continuous. Similarly, if $f : (X, \mathcal{T}) \to (Y, \mathcal{S}_{trivial})$ then f is always continuous.*

Proposition 3.50. *If A is a compact set in $(X, \mathcal{T}_1)$ and $\mathcal{T}_1 \subset \mathcal{T}_2$, then A is compact in $(X, \mathcal{T}_2)$.*

Example 3.51. *Any set A in $(X, \mathcal{T}_{discrete})$, is compact.*

Question 3.52. *Now consider the identity map, $I : X \to Y$. If $I : (X, \mathcal{T}) \to (X, \mathcal{T})$, then clearly $I^{-1}(G) = G$. But what happens for other topologies?*

Proposition 3.53. *1. The identity function $I : (X, \mathcal{T}_1) \to (X, \mathcal{T}_2)$ is continuous if and only if $\mathcal{T}_1 \supset \mathcal{T}_2$.*

2. $I^{-1} : (X, \mathcal{T}_2) \to (X, \mathcal{T}_1)$ is continuous if and only if $\mathcal{T}_1 \subset \mathcal{T}_2$.

3. $I : (X, \mathcal{T}_1) \to (X, \mathcal{T}_2)$ is a homeomorphism if and only if $\mathcal{T}_1 = \mathcal{T}_2$.

Theorem 3.54. *Two metric spaces (X, d_1) and (X, d_2) are the same topological space if and only if they have the same collection of convergent sequences with the same limits. That is*

$$x_n \to x \text{ in } \mathcal{T}_{d_1} \iff x_n \to x \text{ in } \mathcal{T}_{d_2}$$

Proof. ($\Rightarrow$) Trivial.
($\Leftarrow$) We prove this by demonstrating the identity map is a homeomorphism. Observe, for any sequence $x_n \to x$ in (X, d_1), then by assumption

$$I(x_n) \to I(x)$$

in (X, d_2). So $I : (X, d_1) \to (X, d_2)$ is continuous. Thus $\mathcal{T}_{d_1} \supset \mathcal{T}_{d_2}$. Conversely, $\mathcal{T}_{d_1} \subset \mathcal{T}_{d_2}$. Therefore by the previous proposition. $\mathcal{T}_{d_1} = \mathcal{T}_{d_2}$. ∎

Example 3.55. *(X, d) and $(X, \lambda \cdot d)$ where $\lambda > 0$ have the same metric topology.*

Example 3.56. *$(\mathcal{R}^n, d_p)$ for any $1 \le p \le \infty$ induce the same topology:*

$$d_p(\vec{x}_m, \vec{x}) = \sqrt[p]{\sum_{i=1}^{n} |x_{m_i} - x_i|^p} \to 0$$

Midterm Announcement: Midterm will cover Chapters 1,2, and 4. There will be one question from each. And then there will be a fourth mystery question.

3.5 Product Topologies

Definition 3.57. *Given X, Y, then the product set:*

$$X \times Y := \{(x, y) : x \in X, y \in Y\}$$

Example 3.58. *Given $X_i, i \in \mathbb{N}$, then*

$$\prod_{i=1}^{n} X_i = \{(x_1, x_2, \ldots) : x_i \in X_i, i \in \mathbb{N}\} = \{x : \mathbb{N} \to \bigcup_{i=1}^{\infty} X_i : x(n) \in X_n\}$$

Example 3.59. *Given $X_\alpha, \alpha \in \Gamma$, then*

$$\prod_{\alpha \in \Gamma} X_\alpha = \{x : \Gamma \to \bigcup_{\alpha \in \Gamma} X_\alpha : x(\alpha) \in X_\alpha\}$$

Definition 3.60. *For each $\alpha \in \Gamma$. Then we can define the projection map $P_\alpha : \prod_{\alpha \in \Gamma} X_\alpha \to X_\alpha$ by*

$$x \overbrace{\to}^{P_\alpha} x(\alpha)$$

Now, assume each X_α has a topology $\mathcal{T}_\alpha$. We want to define a topology on $\prod_{\alpha \in \Gamma} X_\alpha$ such that each projection map P_α is continuous. That is, we want

$$P_\alpha^{-1}(U_\alpha) \text{ open whenever } U_\alpha \in \mathcal{T}_\alpha$$

Consider the necessary sub-base for this to be true:

$$\phi := \{P_\alpha^{-1}(U_\alpha) : U_\alpha \in \mathcal{T}_\alpha.\alpha \in \Gamma\}$$

Note 3.61. *Since $X_\alpha \in \mathcal{T}_\alpha$, then we have*

$$P_\alpha^{-1}(X_\alpha) = \prod_{\alpha \in \Gamma} X_\alpha$$

We must construct a unique topology having ϕ as its subbase. We shall call this topology the product topology. Moreover, our base $\mathcal{B}$ must satisfy

$$\forall U \in \mathcal{B}, U = P_{\beta_1}^{-1}(U_{\beta_1}) \cap P_{\beta_2}^{-1}(U_{\beta_2}) \cap \ldots \cap P_{\beta_n}^{-1}(U_{\beta_n}) = \bigcap_{i=1}^{n} P_{\beta_i}^{-1}(U_{\beta_i}) \text{ with } U_{\beta_i} \in \mathcal{T}_{\beta_i}$$

Theorem 3.62. *Let Y be a topological space. Then $f : Y \to \prod_{\alpha \in \Gamma} X_\alpha$ continuous if and only if*

$$P_\alpha \circ f : Y \to X_\alpha$$

is continuous for all $\alpha \in \Gamma$.

Theorem 3.63. *Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in the product space $\prod_{\alpha \in \Gamma} X_\alpha$. Then $\{x_n\}_{n=1}^{\infty} \to x$ in $\prod_{\alpha \in \Gamma} X_\alpha$ if and only if for all α,*

$$\{P_\alpha\}_{n=1}^{\infty} \to P_\alpha(x) \text{ in } X_\alpha$$

Example 3.64. *Let X be any non-empty set, and Y be a topological space. Consider all maps*

$$\{f : X \to Y\} = \prod_{x \in X} Y = Y^X$$

Then $\{f_n\}_{n=1}^{\infty}$ converges to f in $\prod_{x \in X} Y$ with respect to the product topology if and only if

$$\forall x \in X, \{P_x(f_n)\} \to P_x(f) \text{ in } Y$$

$$\iff \forall x \in X, \{f_n(x)\} \to f(x)$$

$$f_n \to f \text{ point-wise}$$

Conclusion: The product topology $\prod_{x \in X} Y = \mathcal{T}_{\text{Pointwise}}$!!!

Theorem 3.65 (Tychonoff's Theorem)**.** *Suppose X_α is compact, then for all $\alpha \in \Gamma$, the the product space*

$$\prod_{\alpha \in \Gamma} X_\alpha$$

is also compact.

4 Banach Spaces

4.1 Defining Banach Spaces

Definition 4.1. *A complete, normed linear space $(X, \|\cdot\|)$ is called a Banach space.*

Example 4.2. $(\mathbb{R}^n, \|\cdot\|_p)$ *with* $1 \leq p \leq \infty$*. To show this is a Banach space, we need to prove the triangle-inequality. This is done using Hölder's Inequality:*

$$\sum_{i=1}^{n} |x_i||y_i| \leq \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |y_i|^q \right)^{\frac{1}{q}}$$

where

$$\frac{1}{p} + \frac{1}{q} = 1$$

In particular, when $p = 2$*, then you arrive at the Cauchy-Schwarz Inequality:*

$$\sum_{i=1}^{n} |x_i||y_i| \leq \sqrt{\sum_{i=1}^{n} |x_i|^2} \sqrt{\sum_{i=1}^{n} |y_i|^2}$$

Lemma 4.3. *For* $a \geq 0, b \geq 0$*, then*

$$a^x b^{1-x} \leq xa + (1-x)b$$

for all $x \in [0, 1]$

Proof. Case: $b = 0 \implies$ conclusion immediate!
Case: $b \neq 0$. Notice $f(x) = (\frac{a}{b})^x$ is convex:

$$\begin{aligned} & f(x) \leq xf(1) + (1-x)f(0) \\ \iff & (\frac{a}{b})^x \leq x(\frac{a}{b}) + (1-x) \\ \iff & a^x b^{1-x} \leq ax + b(1-x) \end{aligned}$$

■

Theorem 4.4. *For* $x, y \in \mathbb{R}^n$*, with* $p \geq 1$*, then*

$$\sum_{i=1}^{n} |x_i||y_i| \leq \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |y_i|^q \right)^{\frac{1}{q}}$$

where $\frac{1}{p} + \frac{1}{q} = 1$.

Proof.

$$\begin{aligned} & \sum_{i=1}^{n} |x_i||y_i| \leq \sum_{i=1}^{n} \frac{1}{p}|x_i|^p + \frac{1}{q}|y_i|^q & \text{By previous lemma} \\ \iff & \sum_{i=1}^{n} \frac{|x_i|}{\|x\|_p} \||y_i|\| \, \|y\|_p \leq \sum_{i=1}^{n} \frac{1}{p} \left(\frac{|x_i|}{\|x\|} \right)^p + \frac{1}{q} \left(\frac{|y_i|}{\|y\|} \right)^q = \frac{1}{p} + \frac{1}{q} = 1 \end{aligned}$$

■

Theorem 4.5. *For $x, y \in \mathbb{R}^n$, with $p \geq 1$, then*

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}}$$

Proof. It is clear for $p = 1, \infty$. So we may assume $1 < p < \infty$. Observe:

$$\begin{aligned}
\|x+y\|_p^p &= \sum_{i=1}^{n} |x_i + y_i|^p \\
&= \sum_{i=1}^{n} |x_i + y_i||x_i + y_i|^{p-1} \\
&\leq \sum_{i=1}^{n} |x_i||x_i + y_i|^{p-1} + \sum_{i=1}^{n} |y_i||x_i + y_i|^{p-1} \\
&\leq \|x\|_p \left(\sum_{i=1}^{n} |x_i + y_i|^{q(p-1)}\right)^{\frac{1}{q}} + \|y\|_p \left(\sum_{i=1}^{n} |x_i + y_i|^{q(p-1)}\right)^{\frac{1}{q}} && \text{Hölder's Inequality} \\
&= \left(\|x\|_p + \|y\|_p\right) \left(\sum_{i=1}^{n} |x_i + y_i|^{q(p-1)}\right)^{\frac{1}{q}} \\
&= \left(\|x\|_p + \|y\|_p\right) \|x+y\|_p^{\frac{p}{q}} \\
&= \left(\|x\|_p + \|y\|_p\right) \|x+y\|_p^{p-1}
\end{aligned}$$

$$\implies \|x+y\|_p \leq \|x\|_p + \|y\|_p$$

■

Naturally, we would like to extend these theorems beyond a finite n. This leads us to define sequence spaces:

Example 4.6. *$\ell^p = \ell^p(\mathbb{N}) :=$ the set of all sequences $x = \{x_n\}_{n=1}^{\infty}$ with*

$$\left(\sum_{i=1}^{\infty} |x_i|^p\right)^{\frac{1}{p}} < \infty$$

We can then set $\|x\|_p = \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{\frac{1}{p}}, 1 \leq p < \infty$. The space $(\ell^p, \|\cdot\|_p)$ is a Banach Space.

Example 4.7. *$\ell^\infty :=$ the set of all bounded sequences $x = \{x_n\}_{n=1}^{\infty}$ with the sup-norm*

$$\|x\|_\infty = \sup_i |x_i| < \infty$$

$(\ell^\infty, \|\cdot\|_\infty)$ is a Banach Space.

Theorem 4.8. *Let X be a separable Banach Space. Then there is an isometric embedding from X to ℓ^∞.*

Proof. Do on your own time. ■

This theorem tells us that instead of having to work in a complicated space X, instead, you can study the properties of your space by using a subset of the well known set of bounded sequences.

Example 4.9. *Let X be a compact metric space. Then $(C(X), \|\cdot\|_\infty)$ is Banach.*

Example 4.10. *Let $k \geq 1$. Then $C^k(X) :=$ the space of k-times continuously differentiable functions on $[a,b]$. Then $(C^k([a,b]), \|\cdot\|_\infty)$ is sadly not a Banach Space. We need to define a norm that will give us this property. Define*

$$\|f\|_{C^k} = \|f\|_\infty + \|f'\|_\infty + \ldots + \left\|f^{(k)}\right\|_\infty$$

Then $(C^k([a,b]), \|\cdot\|_{C^k})$ is Banach!

Example 4.11. *$(C([a,b), \|\cdot\|_p)$ is a normed linear space with $1 \leq p < \infty$, where*

$$\|f\|_p := \left(\int_a^b |f(x)|^p dx\right)^{\frac{1}{p}}$$

But $(C([a,b), \|\cdot\|_p)$ is not Complete!

Of course we would like to have completion on the last example. We can instead define a space that will suffice:

Example 4.12. *The completion of $(C[a,b], \|\cdot\|_p)$ is*

$$L^p([a,b]) := \{f : [a,b] \to \mathbb{R} : \left(\int_a^b |f(x)|^p\right)^{\frac{1}{p}} < \infty\}$$

Then $(L^p([a,b]), \|\cdot\|_p)$ is Banach for $1 \leq p < \infty$.

Now we recall some notions from linear algebra for finite-dimensional spaces:

Definition 4.13. *The dimension of a linear space is equal to the number of elements that form a (linear) basis for the space.*

Theorem 4.14. *Every linear space X has a basis which is a maximal, linearly independent subset B, called an algebraic basis or Hamel Basis.*

So for every $x \in X$, there exists an n such that

$$e_1, e_2, \ldots, e_n \in B$$

and

$$c_1, c_2, \ldots, c_n \in \mathbb{R}$$

such that x can be writeen as a unique linear combination of basis elements:

$$x = c_1 e_1 + c_2 e_2 + \ldots + c_n e_n$$

This leaves something to be desired in general spaces: Every vector is a finite combination of basis vectors. Of courses, this can't happen when B is too large, i.e. $C([0,1])$. But we still want to extend these ideas into the realm of Banach spaces.

Definition 4.15. *Let X be a separable Banach space. A Schauder Basis is a sequence $\{x_n\}_{n=1}^\infty$ such that for all $x \in X$, there exists a unique sequence $\{c_n\}_{n=1}^\infty \subset \mathbb{R}$ such that*

$$x = \sum_{i=1}^\infty c_n x_n$$

Remark 4.16. *Doesn't always exists for Banach Spaces.*

Example 4.17. *Consider $C([0,1])$ equipped with $\|\cdot\|_\infty$ has a basis composed of hill functions* ***FINISH***

4.2 Mappings Between Normed Linear Spaces

Definition 4.18. *Let X, Y be linear spaces. Then $T : X \to Y$ is linear provided*

$$T(ax + by) = aT(x) + bT(y)$$

for any $x, y \in X, a, b \in \mathbb{F}$.

Definition 4.19.

$$\mathcal{L}(X,Y) := \{T : X \to Y : T \text{ is linear } \}$$

For this new space, T may not be finite dimensional, which brings us concerns of continuity. We will approach this by considering boundedness first (and then learn that boundedness and continuity are equivalent for linear maps).

Definition 4.20. *Let X, Y be normed, linear spaces. A linear map $T : X \to Y$ is bounded if there exists $M \geq 0$ such that*

$$\|Tx\|_Y \leq M \|x\|_X$$

Definition 4.21.

$$\mathcal{B}(X,Y) := \{T : X \to Y : T \text{ is a bounded, linear map from } X \text{ to } Y\}$$

Note 4.22. *In the case where $X = Y$, we will simply write $\mathcal{L}(X) := \mathcal{L}(X,X)$ and $B(X) := B(X,X)$.*

Definition 4.23. *Given a space X, $X^* = B(X,\mathbb{R})$ is called the (topological) dual space of X.*

This space is very handy when trying to analyze X. Now we study bounded, linear maps:

Definition 4.24. *Given the space $B(X,Y)$ we can define the operator norm*

$$\|T\| = \inf\{M : \forall x \in X,\ \|Tx\| \leq M \|x\|\}$$

Clearly, if $x \neq 0$

$$\frac{\|Tx\|}{\|x\|} \leq M \implies \|T\| := \sup_{\|x\| \neq 0} \frac{\|Tx\|}{\|x\|}$$

Note 4.25. $|\lambda| \, \|T\| = \|\lambda Tx\|$. *So*

$$\frac{\|Tx\|}{\|x\|} = \left\| T\left(\frac{x}{\|x\|}\right) \right\|$$

which allows us to refine our definition of the operator norm:

$$\|T\| = \sup_{\|x\|=1} \|Tx\| = \sup_{0<\|x\|<1} \frac{\|Tx\|}{\|x\|}$$

Note 4.26. *For all $x \in X$*

$$\|Tx\| \leq \|T\| \cdot \|x\|$$

Example 4.27. *The Identity map is bounded and linear*

$$I : X \to X$$

with norm $\|I\| = 1$

Clearly, boundedness is boring. What may be surprising is an unbounded example:

Example 4.28. *Let $X = C^\infty([0,1])$ and consider*

$$T = \frac{d}{dx} : X \to X$$

This is, in fact, unbounded, for any norm on X. Observe, let $f_\lambda(x) = e^{\lambda x} \in X$. Then

$$Tf_\lambda = \lambda f_\lambda(x)$$

So

$$\|Tf_\lambda\| = \|\lambda f_\lambda\| = |\lambda| \, \|f_\lambda\|$$

Since $|\lambda|$ is arbitrary, we can take it to be as large as we want, do

$$\|Tx\|_Y \not\leq M \, \|X\|_X$$

Example 4.29. *If $B(\mathbb{R}) = \{A : \mathbb{R} \to \mathbb{R} | Ax = ax, a \in \mathbb{R}\} \cong \mathbb{R}$. Here*

$$\|A\| = |a|$$

Example 4.30. *$X \in \mathbb{R}^n$, $Y = \mathbb{R}^m$. Then any linear transformation $T : X \to Y$ is represented by a matrix*

$$A := (a_{i,j})_{m \times n}$$

such that $T(x) = A \cdot x$.

Question 4.31. *What is $\|A\|$?*

This norm is different depending on the topology used.

Example 4.32. *Consider X, Y equipped with the Euclidean norm $\|\cdot\|_2$. Then*

$$\|x\|_2 = \sqrt{x_1^2 + \ldots + x_n^2} = \langle x, x \rangle$$

Then

$$\|A\|_2^2 = \sup_{\|x\|_2 = 1} \|Ax\|_2^2$$

Using the method of Lagrange Multipliers,

$$\begin{aligned} f(x, \lambda) &= \langle Ax, Ax \rangle + \lambda(1 - \langle x, x \rangle) \\ &= \langle x, A^T Ax \rangle - \lambda \langle x, x \rangle + \lambda \\ &= \langle x, (A^T A - \lambda I)x \rangle + \lambda \end{aligned}$$

The critical points of this equation (setting $f(x, \lambda) = 0$) are characterized by

$$(A^T A - \lambda I) = 0$$

Which is equivalent to determining the eigenvalues of $A^T A$. So

$$0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$$

each of which have an eigenvector $x_i, 1 \leq i \leq n$. Then

$$\implies f(x_i, \lambda_i) = \lambda_i$$

$$\implies \|A\|_2^2 = \sup_i \lambda_i = \text{ max eigenvalue of } A^T A$$

$$\implies \|A\|_2 = \sqrt{\text{ max eigenvalue of } A^T A}$$

Example 4.33. *Same as before, but replace* $\|\cdot\|_2$ *with* $\|\cdot\|_\infty = \max_i\{|x_i|\}$. *Then*

$$\|A\|_\infty = \sup_{\|x\|_\infty \le 1} \|Ax\|_\infty$$

Let $y = Ax$, *then*

$$y_i = \sum_{j=1}^n a_{i,j} x_j$$

$$\implies |y_i| = \left|\sum_{j=1}^n a_{i,j} x_j\right| \le \sum_{j=1}^n |a_{i,j}||x_j| \le \|x\| \sum_{j=1}^n |a_{i,j}| \le \sum_{j=1}^n |a_{i,j}|$$

$$\implies \|y\|_\infty = \max_i\{|y_i\} \le \max_i \sum_{j=1}^n |a_{i,j}| = \textit{ maximum row sum}$$

Lemma 4.34. $\|A\|_\infty =$ *maximum row sum*

Proof. Case: If $A = 0$, then the conclusion is immediate.
Case: Suppose $A \neq 0$ and suppose the max row sum is given by the sum

$$\sum_{j=1}^n |a_{i_0,j}|$$

for some $1 \le i_0 \le m$. Let $x_j = sign(a_{i_0,j})$ and set

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Clearly, $\|x\|_\infty = 1$. Then

$$\|Ax\|_\infty = \sum_{j=1}^n |a_{i_0,j}| = \text{ maximum row sum}$$

$$\implies \|A\|_\infty = \text{ maximum row sum}$$

■

Theorem 4.35. *Let* $T : X \to Y$ *be a linear map between normed linear spaces. Then the following statements are equivalent:*

1. $\|T\| = \sup_{\|x\|=1} \|Tx\| < \infty$.
2. $T \in B(X, Y)$, *that is,* T *is bounded.*
3. T *is uniformly continuous on* X.
4. T *is continuous at* $0 \in X$.

Proof.
- (1) $\implies$ (2). That is, for all $x \in X$, $\|Tx\| \le \|T\| \|x\|$. So taking $M = \|T\|$ proves boundedness.
- (2) $\implies$ (3).
  Claim: $\|T\| = 0$, then $\|Tx\| = 0 \implies Tx = 0 \implies T = 0 \implies$ T is uniformly continuous.
  Claim: $\|T\| > 0$, then for all $\epsilon > 0$, let $\delta = \frac{\epsilon}{\|T\|}$. Then for all $x, y \in X$ with $\|x - y\| < \delta$, we have

$$\|Tx - Ty\| = \|T(x - y)\| \le \|T\| \|x - y\| \le \|T\| \cdot \delta \le \epsilon$$

- (3) $\implies$ (4). Trivial.
- (4) $\implies$ (1). Suppose T is continuous at $0 \in X$. Then for $\epsilon = 1$, there exists δ such that

$$\|x - 0\| \leq \delta \implies \|Tx\| \leq \epsilon = 1$$

Therefore, for every $x \in X$ with $\|x\| = 1$, then

$$\|Tx\| = \left\| \frac{1}{\delta} T(\delta x) \right\| = \frac{1}{\delta} \|T(\delta x)\| \leq \frac{1}{\delta}$$

Therefore,

$$\|T\| \leq \frac{1}{\delta} < \infty$$

■

Theorem 4.36. *Let X and Y be normed, linear spaces. Then $B(X,Y)$ is also a normed, linear space equipped with the operator norm*

$$\|T\|_{operator} := \sup_{\|x\|=1} \|Tx\|$$

for all $T \in B(X,Y)$.

Proof. 1. First, we need to prove that $B(X,Y)$ is a linear subspace of $\mathcal{L}(X,Y)$.

- For all $S, T \in B(X,Y)$, we can take $x \in X$ such that $\|x\| = 1$. Then

$$\begin{aligned} \|(S+T)x\| &= \|Sx + Tx\| \\ &\leq \|Sx\| + \|Tx\| \\ &\leq \|S\| + \|T\| \end{aligned}$$

Therefore, $S + T \in B(X,Y)$ and $\|S + T\| \leq \|S\| + \|T\|$.

- For all $\lambda \in \mathbb{R}$ (or $\mathbb{C}$), then for $T \in B(X,Y)$, we see

$$\begin{aligned} \|\lambda T\| &= \sup_{\|x\|=1} \|\lambda T(x)\| \\ &= \sup_{\|x\|=1} |\lambda| \|Tx\| \\ &= |\lambda| \|T\| < \infty \end{aligned}$$

Therefore $\lambda T \in B(X,Y)$ and $\|\lambda T\| = |\lambda| \|T\|$

These two properties demonstrate $B(X,Y)$ is a linear space.

2. Next we need to show that the operator norm is satisfactory as a norm on $B(X,Y)$.

- $\|T\| \geq 0$ is inherited from its definition.
- $\|0\| = 0$
- If $\|T\| = 0 \iff T = 0$

Therefore $\|T\|_{\text{operator}}$ is a norm on $B(X,Y)$.

■

Proposition 4.37. *If $S \in B(X,Y), T \in B(Y,Z)$, then*

$$T \circ S \in B(X,Z) \text{ with } \|TS\| \leq \|T\| \cdot \|S\|$$

Remark: This shows that $B(X) = B(X, X)$ is a normed algebra.

Proof. Observe:

$$\|TSx\| \leq \|T\| \, \|Sx\| \leq \|T\| \, \|S\| \, \|x\| \implies \|TS\| \leq \|T\| \, \|S\|$$

■

Theorem 4.38. *Let X be a normed, linear space, and Y be a Banach space. Then $B(X, Y)$ is a Banach space.*
In particular for any normed, linear space X, the dual space $X^* = B(X, \mathbb{R})$ is a Banach space. Also, if X is Banach, then $B(X)$ is also Banach.

Proof. Let $\{T_n\}_{n=1}^{\infty}$ be a Cauchy sequence in $B(X, Y)$. Then for all $\epsilon > 0, \exists N$ such that $\forall m, n > N$

$$\|T_n - T_m\| < \epsilon$$

1. We need to show that the sequence $\{T_n(x)\}_{n=1}^{\infty}$ is Cauchy. Observe, for any $x \in X$

   $$\|T_n(x) - T_m(x)\| \leq \|T_n - T_m\| \, \|x\| \leq \epsilon \cdot \|x\|$$

   Therefore, $\{T_n(x)\}_{n=1}^{\infty}$ is Cauchy.

2. Is there a candidate for convergence? Yes, we define

   $$T(x) = \lim_{n \to \infty} T_n(x) \in Y$$

   This is well defined because Y is Banach and the Cauchy sequence $\{T_n(x)\}_{n=1}^{\infty}$ converges pointwise.

3. Linearity: Since T_n is linear, then

   $$T(ax + by) = \lim_{n \to \infty} T_n(ax + by) = \lim_{n \to \infty} aT_n(x) + bT_n(y) = a \lim_{n \to \infty} T_n(x) + b \lim_{n \to \infty} T_n(y) = aT(x) + bT(y)$$

4. Boundedness: Since $\{T_n\}_{n=1}^{\infty}$ is Cauchy, then it is also bounded. That is, there exists $M > 0$ such that

   $$\|T_n\| \leq M$$

   Thus, for all $x \in X$,

   $$\|T_n x\| \leq \|T_n\| \cdot \|x\| \leq M \, \|x\|$$
   $$\|Tx\| = \left\| \lim_{n \to \infty} T_n(x) \right\| \leq M \, \|x\|$$
   $$\implies \|T\| \leq M \text{ and } T \in B(X, Y)$$

5. Lastly, we need to show uniform convergence to T. Let $m \to \infty$, then

   $$\|T_n x - Tx\| < \epsilon \, \|x\| \implies \|T_n - T\| \leq \epsilon \implies T_n \to T \text{ in } B(X, Y)$$

Therefore, $B(X, Y)$ is complete. Hence, it is Banach. ■

4.3 Equivalent Norms on Finite Dimensional Spaces

Definition 4.39. *Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on a linear space X. We say $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$ if*

$$\|x_n\|_2 \to 0 \implies \|x_n\|_1 \to 0$$

Lemma 4.40. *$\|\cdot\|_2$ is stronger than $\|\cdot\|_1$ if and only if there exists $C >$ such that*

$$\|x\|_1 \leq C \|x\|_2$$

for every $x \in X$.

Proof. Homework. ∎

Example 4.41. *The identity $I : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$ is bounded as a result of the lemma.*

Definition 4.42. *Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent if there exists $c_1 > 0$ and $c_2 > 0$ such that*

$$c_1 \|x\|_1 \leq \|x\|_2 \leq c_2 \|x\|_1$$

for all $x \in X$.

Remark 4.43. *This defines an equivalence relation on norms. Say,*

$$\|\cdot\|_1 \sim \|\cdot\|_2$$

if and only if there exists a $C > 0$ such that

$$\frac{1}{C} \|x\|_1 \leq \|x\|_2 \leq C \|x\|_1$$

Theorem 4.44. *Two norms on a linear space generate the same topology if and only if the norms are equivalent.*

Proof. Let $I : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$ be the identity map. Recall, that

$$\begin{aligned} \text{topologies are the same} &\iff \text{both } I \text{ and } I^{-1} \text{ are continuous} \\ &\iff \text{both } I \text{ and } I^{-1} \text{ are bounded} \\ &\iff \text{norms are equivalent} \end{aligned}$$

∎

Question 4.45. *Let X be a finite-dimensional linear space. How many norms can we define?*

Theorem 4.46. *Every two norms on a finite-dimensional linear space are equivalent. To answer the question above, there is one unique norm up to equivalence.*

Proof. 1. Find an norm on an arbitrary finite-dimensional space.
Let $n := dim(X)$. Then we can write $\{e_1, \ldots, e_n\}$ as the basis of X. Define

$$f : \mathbb{R}^n \to X$$

$$x = (x_1, \ldots, x_n) \to \sum_{i=1}^{n} x_i e_i$$

Then, f is an isomorphism. On $\mathbb{R}^n$, we have the standard euclidean norm,

$$|x| = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Using f, we can define a norm on X by setting

$$\begin{aligned}\|f(x)\|_{\mathbb{R}^n} &= \left\|\sum_{i=1}^{n} x_i e_i\right\|_{\mathbb{R}^n} \\ &:= |x| = \sqrt{\sum_{i=1}^{n} x_i^2}\end{aligned}$$

2. Show that any norm $\|\cdot\|$ on X is equivalent to $\|\cdot\|_{\mathbb{R}^n}$
Indeed, let $p : \mathbb{R}^n \to \mathbb{R}$ to be

$$x \to \|f(x)\|$$

Claim: p is Lipschitz continuous. That is, for all $x, y \in \mathbb{R}^n$,

$$\begin{aligned}\|p(x) - p(y)\| &= \left\|\left\|\sum_{i=1}^{n} x_i e_i\right\| - \left\|\sum_{i=1}^{n} y_i e_i\right\|\right\| \\ &\leq \left\|\sum_i (x_i - y_i) e_i\right\| \\ &\leq \sum_i |x_i - y_i| \, \|e_i\| \\ &\leq \sqrt{\sum_i |x_i - y_i|^2} \sqrt{\sum_i \|e_i\|^2} && \text{Cauchy-Schwarz} \\ &= |x - y| \sqrt{\sum_i \|e_i\|^2}\end{aligned}$$

Therefore, p is Lipschitz continuous!.

Therefore, p achieves its extreme values on the unit sphere,

$$B_1(0) = \{x : \mathbb{R}^n : |x| = 1\}$$

which is compact. So for some $C \geq c \geq 0$

$$c \leq p(x) \leq C$$

for all $x \in B_1(0)$. (Notice that $c > 0$, because otherwise $p(x^*) = 0$ for some $x^* \in \mathbb{R}^n$ with $|x^*| = 1$, which implies

$$\|f(x^*)\| = 0 \implies f(x^*) = 0 \implies x^* = 0$$

but this contradicts what we said previously that $|x^*| = 1$.) Therefore,

$$c \leq \|f(x)\| \leq C$$

for all $x \in \mathbb{R}, |x| = 1$. To generalize this, when $x \neq 0$, and by the fact that f is linear, then

$$\|f(x)\| = \left\| \|x\| \, f\left(\frac{x}{\|x\|}\right)\right\| = |x| \left\|f\left(\frac{x}{\|x\|}\right)\right\|$$

Since we proved the equivalence chain for unit vectors in X, we see that

$$\begin{aligned} c|x| \leq \|f(x)\| \leq C|x| \\ \iff c\,\|f(x)\|_{\mathbb{R}^n} \leq \|f(x)\| \leq C\,\|f(x)\|_{\mathbb{R}^n}\end{aligned}$$

Therefore, $\|\cdot\|$ is equivalent to $\|\cdot\|_{\mathbb{R}^n}$.

■

Corollary 4.46.1. *Every linear map on a finite dimensional, linear space is bounded. Moreover,*

$$\mathcal{L}(X,Y) = B(X,Y)$$

whenever $dim(X) < \infty$.

Proof. Suppose $A \in \mathcal{L}(X,Y)$ and $dim(X) = n < \infty$. Then we can write $\{e_1, \dots, e_n\}$ as the basis of X. Then for every $x \in X$ can be written

$$x = \sum_{i=1}^{n} x_i e_i$$

So

$$\|Ax\| = |\sum_{i=1}^{n} x_i Ae_i| \leq \sum_{i=1}^{n} |x_i| \, \|Ae_i\| \leq \max_{1 \leq i \leq n} \|Ae_i\| \sum_{i=1}^{n} |x_i| \leq C \max_{1 \leq i \leq n} \|Ae_i\| \, \|x\|$$

for some $C > 0$ by the equivalence of norms on finite dimensional spaces. Therefore, $A \in B(X,Y) \implies \mathcal{L}(X,Y) \subset B(X,Y) \implies \mathcal{L}(X,Y) = B(X,Y)$ ■

Corollary 4.46.2. *Every finite dimensional normed linear space X is a Banach space.*

Proof. Because any norm on X is equivalent to $\|\cdot\|_{\mathbb{R}^n}$, which is complete. ■

Corollary 4.46.3. *Every finite dimensional linear subspace A of a normed linear space X (not necessarily finite-dimensional), is closed.*

Proof. Because every finite dimensional normed linear space is complete, and A a complete subset requires A to be closed. ■

Remark 4.47. *If $dim(A) = \infty$, A may fail to be closed.*

Example 4.48. $X = \ell^\infty :=$ \{ *bounded sequence* $\{x_n\}_{n=1}^{\infty}$\} *with* $\|\cdot\|_\infty$. *Then for all n, let*

$$A_n := \{(x_1, x_2, \dots, x_n, 0, 0, \dots) \in X\}$$

Clearly A_n is a closed, linear subspace of X. So we can have the chain:

$$A_1 \subset A_2 \subset \dots \subset A_n \subset \dots \subset X$$

Let

$$A := \bigcup_{n=1}^{\infty} A_n$$

Then A is also a linear subspace of A, but not closed. In particular, for every $n \in \mathbb{N}$:

$$a_n = (1, \frac{1}{2}, \dots, \frac{1}{n}, 0, 0, \dots) \in A_n \subset A$$

$$a_n \to \left\{\frac{1}{n}\right\}_{n=1}^{\infty} \notin A$$

Theorem 4.49 (Bounded Linear Transformation)**.** *Let X be a normed, linear space, and Y a Banach space. If M is a dense linear subset of X and $T : M \subset X \to Y$ is a bounded linear map, then there exists a unique linear map $\overline{T} : X \to Y$ such that $\overline{T}x = Tx$ for all $x \in M$. Moreover,*

$$\|\overline{T}\| = \|T\|$$

Proof. Recall we already did something similar to this in the homework when proving the Lipschitz extension of a function defined on a dense set. Now recall, for a Lipschitz map, there is a constant M such that

$$\|f(x) - f(y)\| \leq M \|x - y\|$$

Similarly, in this context, we can leverage linearity to have a similar result:

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

So for each bounded linear map $S : X \to Y$, it's also a Lipschitz map with $Lip(S) = \|S\|$. By the Lipschitz extension theorem, from the previous homework, there exists a unique Lipschitz map $\overline{T} : X \to Y$ such that $\overline{T}|_M = T$ and $Lip(\overline{T}) = Lip(T)$. By this extension theorem, $\overline{T}$ is defined by: for any $x \in X$

$$\overline{T}(x) = \lim_{n\to\infty} Tx_n$$

for any $x_n \to x, x_n \in M$. Since T is linear, $\overline{T}$ must also inherit this linearity from T under this definition. Lastly, we need to show this mapping is bounded. This

$$\left\|\overline{T}\right\| = Lip(\overline{T}) = Lip(T) = \|T\| < \infty$$

Therefore, $\overline{T} \in B(X, Y)$. ■

Remark 4.50. *If $M \subset X$ is not dense, then one can extend $T : M \to Y$ to $\overline{T} : M \to Y$, where*

$$\overline{M} = closure(M)$$

As a result of $T \in B(X, Y)$, where Y is Banach, then we may assume that the domain of T is closed.

4.4 The Kernel and Range of a Bounded Linear Map

Let X and Y be linear spaces, and $T \in \mathcal{L}(X, Y)$.

Definition 4.51.

$$Ker(T) := \{x \in X : Tx = 0\}$$

It is a subspace of X, and also called the null space of T.

Definition 4.52.

$$Range(T) := \{y \in Y : Tx = y \text{ for some } x \in X\}$$

It is a subspace of Y.

Note 4.53. *T is one-to-one if and only if $Ker(T) = \{0\}$.*
T is onto if and only if $range(T) = Y$.

Proposition 4.54. *If X and Y are normed linear spaces, and $T \in B(X, Y)$, then*

$$Ker(T) \text{ is closed}$$

Proof. Simple Proof: T bounded $\implies$ T continuous $\implies$ $Ker(T) = T^{-1}(\{0\})$ is closed.

Undergraduate Linear Algebra Proof: For all $x_n \in Ker(T)$ with $x_n \to x$ in X. By the continuity of T,

$$Tx = \lim_{n\to\infty} Tx_n = \lim_{n\to\infty} 0 = 0$$

Therefore, $x \in Ker(T)$ meaning $Ker(T)$ is closed. ■

Question 4.55. *Recall $T \in \mathcal{X}$ with $dim(X) < \infty$. Then*

$$Ker(T) = 0 \iff Range(T) = X$$

Can this notion be extended for ∞ dimensional X? No! Consider the example:

Example 4.56. *We consider two operators on ℓ^∞ (the set of bounded sequences), by defining:*

$$\textit{right-shift}: S : \ell^\infty \to \ell^\infty$$

$$S(x_1, x_2, \dots, x_n, \dots,) = (0, x_1, x_2, \dots, x_n, \dots)$$
$$\textit{left-shift}: T : \ell^\infty \to \ell^\infty$$
$$T(x_1, x_2, \dots, x_n, \dots,) = (x_2, x_3, \dots, x_n, \dots)$$

Then

$$\|S\| = \sup_{x \neq 0} \frac{\|Sx\|_\infty}{\|x\|_\infty} = 1$$

Also,

$$\|T\| = \sup_{2 \le i < \infty} |x_i| \le \sup_{1 \le i \le \infty} |x_i| = \|x\|_\infty \implies \|T\| \le 1$$

To find the lower bound, we can just find a sequence that maintains this relationship:

$$\|T(1, 1, 1, \dots)\|_\infty = \|(1, 1, 1, \dots)\|_\infty \implies \|T\| \ge 1$$

Together, these show $\|T\| = 1$.

Now

$$Ker(S) = \{0\} \qquad Range(S) = \{x \in \ell^\infty : x_1 = 0\} \neq \ell^\infty$$

This defeats our original notions in the previous question. That is, S is injective but not onto. Also,

$$Ker(T) = \{x \in \ell^\infty : x_i = 0 \text{ for } i \ge 2\} \qquad Range(T) = \ell^\infty$$

So T is onto but not injective.

Example 4.57. *$X = (C([a, b]), \|\cdot\|_\infty)$ by defining the Volterra Integral Operator by*

$$K : C([a, b]) \to C([a, b])$$

$$f \to Kf(x) := \int_a^x f(t)\ dt$$

That is, Kf is the anti-derivative of f with $Kf(a) = 0$. To calculate its norm:

$$\begin{aligned}
\|Kf\| &= \max_{a \le x \le b} \left| \int_0^x f(t)\ dt \right| \\
&\le \max_{a \le x \le b} \int_0^x \|f\|\ dt \\
&= \max_{a \le x \le b} \|f\| (b - a) \\
&= \|f\| (b - a)
\end{aligned}$$

$$\implies \|K\| \le (b - a)$$

To find a lower bound, we consider a single point at which K achieves this maximum. Specifically:

$$\|K1\| = b - a \implies \|K\| \ge (b - a)$$

These two together show $\|K\| = b - a$. *Now*

$$Ker(K) = \left\{ f \in C([a,b]) : \int_a^x f(t)\ dt \equiv 0 \right\} = \{0\} \implies K \text{ is one-to-one}$$

$$Range(K) = \{ g \in C^1[a,b] : g(a) = 0 \}$$

Question 4.58. *Is* $Range(K)$ *closed? No! You can approximate a cusp with differentiable functions. Therefore,*

$$Range(K) \neq C([a,b]) \implies Kf = g \text{ does not have a solution when } g \notin Range(K) \implies K \text{ not invertible.}$$

4.5 Open Mapping Theorem

Motivation: Very often, we need to solve an equation of the form

$$Tx = y$$

for some given linear operator $T : X \to Y$.

Definition 4.59. *A problem is called* <u>*well-posed*</u> *if*

- *A solution exists*
- *The solution is unique*
- *The solution is "stable", i.e. the solution depends continuously on the data.*

Question 4.60. *What type of operator* T *will result in a well-posed problem?*

1. Existence of a solution $x \in X$ for each $y \in Y$ if T has a right-side inverse operator

$$T_r^{-1} : Y \to X \text{ such that } TT_r^{-1} = I_Y$$

   That is, for every $y \in Y$, we let $x \in T_r^{-1}y$. Then

$$TT_r^{-1}y = y$$

2. At most one solution $x \in X$ requires that T has a left-side inverse operator

$$T_\ell^{-1} : Y \to X \text{ such that } T_\ell^{-1}T = I_X$$

   That is, suppose $Tx = y$ for some x. Then

$$T_\ell^{-1}Tx = x \implies x = T_\ell^{-1}y$$

   Note 4.61. *If* T *has both a left-side inverse* T_ℓ^{-1} *and a right-side inverse* T_r^{-1}, *then*

$$T_\ell^{-1} = T_r^{-1}$$

   Since

$$T_r^{-1} = (T_\ell^{-1}T)T_r^{-1} = T_\ell^{-1}(TT_r^{-1}) = T_\ell^{-1}$$

   In this case, we denote it by T^{-1}. *So* $Tx = y$ *has a unique solution* $\iff$ T^{-1} *exists.*

3. Stability

Definition 4.62. *A solution of $Tx = y$ is* <u>*stable*</u> *if when the original data y is perturbed a little bit, then the corresponding solution x of $Tx = y$ is also changed very little.*

Another way of stating this, is for all $\epsilon > 0, \exists \delta > 0$ such that

$$\|y - \tilde{y}\| < \delta \implies \|x - \tilde{x}\| < \epsilon$$

But clearly this language is identical to that of functional continuity. Moreover,

$$\text{Stability means } T^{-1} : X \to Y \text{ is continuous}$$

Proof. For all open sets $U \subset X$, $(T^{-1})^{-1}(U)$ is open in Y.
$\iff$ for every open set $U \subset X$, TU is open in Y.
$\iff$ T is an open mapping. ■

Definition 4.63. *$T : X \to Y$ is an open mapping if TU is open in Y whenever U is open in X.*

Theorem 4.64 (Open Mapping Theorem)**.** *Let $T : X \to Y$ be a bounded, linear operator between Banach Spaces. If T is onto, then T is an open mapping.*

Remark 4.65. *We may not drop "onto" or "Banach" conditions.*

Example 4.66. *Consider the Volterra Integral Operator:*

$$K : C[a, b] \to C[a, b]$$

$$Kf(x) = \int_a^x f(y)\ dy$$

We explore what happens when you change the norm on $C[a, b]$

1. $(C[a, b], \|\cdot\|_\infty) \to (C[a, b], \|\cdot\|_\infty)$
   Then K is a bounded linear operator and $range(K) := \{g \in C^1[a, b] : g(a) = 0\} \implies K$ is not onto. Also $K^{-1} = \frac{d}{dx}$ but clearly it is not bounded. Therefore, we cannot drop the onto condition.

2. $(C[a, b], \|\cdot\|_\infty) \to range(K) = \{g \in C^1[a, b] : g(a) = 0\}$.
   Then K is now onto, but $(range(K), \|\cdot\|)$ is not a Banach space. This results in K^{-1} not bounded.

3. $(C[a, b], \|\cdot\|_\infty) \to (range(K), \|\cdot\|_{C^1})$.
   Now we have a bounded, linear operator between Banach Spaces. Then $D = K^{-1}$ is continuous! Indeed,

$$\|Dg\|_\infty = \left\|g'(x)\right\|_\infty \leq \|g\|_{C^1} \implies \|D\| \leq 1$$

Corollary 4.66.1 (Banach's Inverse Operator Theorem)**.** *Suppose $T : X \to Y$ is a one-to-one, onto, bounded, linear operator between Banach Spaces. Then $T^{-1} : Y \to X$ is bounded.*

Proof. For every U open subset of X, $(T^{-1})U$ is open in Y because T is an open mapping. So therefore, T^{-1} is continuous $\implies T^{-1}$ is bounded. ■

Theorem 4.67. *Let X, Y be Banach spaces and $T \in B(X, Y)$. Then the following statements are equivalent:*

1. *$\exists\ c > 0$ such that $c\,\|x\| \leq \|Tx\|$ for all $x \in X$*

2. *$Ker(T) = \{0\}$ and $range(T)$ is closed.*

Proof. ($\Leftarrow$) Since $range(T)$ is a closed subspace of a Banach space Y, then $range(T)$ is also a Banach space. Thus $T : X \to range(T)$ is a one-to-one, onto, bounded linear map between Banach spaces. Therefore, by the previous corollary, T is an open mapping with bounded inverse $T^{-1} : range(T) \to X$. That is, there exists $c > 0$ such that

$$\left\|T^{-1}y\right\| \leq c\|y\|$$

for all $y \in range(T)$. Setting $y = Tx$, we have

$$\|x\| \leq C\|Tx\|$$

$$\iff \frac{1}{C}\|x\| \leq \|Tx\|$$

($\Rightarrow$) Suppose there exists $c\|x\| \leq \|Tx\|$ for all $x \in X$. Notice, if $\|Tx\| = 0 \implies \|x\| = 0 \implies x = 0 \implies Ker(T) = \{0\}$ and T is one-to-one. Also, suppose $\{y_n\}$ is a convergent sequence in $range(T)$ with $y_n \to y$ for some $y \in Y$. We want to show that $y \in range(T)$. To show this, observe:

$$y_n = Tx_n$$

for some $x_n \in X$. Then

$$c\|x_n - x_m\| \leq \|Tx_n - Tx_m\| = \|y_n - y_m\|$$

Since y_n is Cauchy, x_n is Cauchy. Since X is complete, $x_n \to x$ for some $x \in X$. By continuity,

$$y_n = Tx_n \to Tx \implies y = Tx$$

■

5 Linear Functionals

5.1 Dual Spaces

Definition 5.1. *Let X be a linear space. Then a linear map $f : X \to \mathbb{R}$ is called a linear functional on X.*

Definition 5.2. *When X is a normed linear space, its (topological) dual space is*

$$X^* := B(X, R)$$

the space of all continuous linear functionals on X.

Note 5.3. *X^* is a Banach space.*

Example 5.4. *Suppose $dim(X) = n < +\infty$. Then $dim(X^*) = n$. We can see this because let:*

$$\{e_1, \dots, e_n\}$$

be a basis of X. Then for each i, we define $w_i : X \to \mathbb{R}$ by:

$$w_i\left(x = \sum_{j=1}^{n} x_j e_j\right) = x_i$$

Clearly, $w_i \in X^*$.

Claim: $\{w_1, w_2, \dots, w_n\}$ forms a basis of X^*, called the dual basis of $\{e_1, \dots, e_n\}$.

Proof.

- We show that any function in X^* can be written as a linear combination of basis elements: Indeed, for every $\phi \in X^*$, and for all $x = x = \sum_{j=1}^{n} x_j e_j$, then

$$\phi(x) = \phi\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{j=1}^{n} x_j \phi(e_j) = \sum_{j=1}^{n} w_j(x)\phi(e_j)$$

  Therefore, $\phi(x) = \sum_{j=1}^{n} w_j(x)\phi(e_j)$

- We show this linear combination is unique:
  Also if

$$\sum_{j=1}^{n} \lambda_j w_j = \vec{0} \in X^*$$

  then it follows for each $1 \le i \le n$:

$$0 = \vec{0}(e_i) = \left(\sum_{j=1}^{n} \lambda_j w_j\right) e_i = \lambda_i$$

Therefore, $\{w_1, \dots, w_n\}$ forms a basis of X^* and $dim(X^*) = n$. ■

Example 5.5. *Consider*

$$\ell^p := \{\{x_i\}_{i=1}^{\infty} : \left(\sum_i |x_i|^p\right)^{\frac{1}{p}} < +\infty\}$$

Then $(\ell^p)^* = \ell^q$ where

$$\frac{1}{p} + \frac{1}{q} = 1$$

Example 5.6.

$$(L^p(\Omega))^* = L^q(\Omega) \text{ for } 1 < p < \infty, \frac{1}{p} + \frac{1}{q} = 1$$

Example 5.7.

$$(C([a,b]), \|\cdot\|_\infty)^* := \text{ the space of Radon measure on } [a,b]$$

Theorem 5.8 (Hahn-Banach Theorem)**.** *Let X be a normed linear space, and Y be a linear subspace of X. Then for any bounded linear function ψ on Y, there exists a bounded linear function ϕ on X such that*

1. *Extension: $\phi(x) = \psi(x)\ \forall\ x \in Y$*
2. *Norm Preservation: $\|\phi\| = \|\psi\|$*

Corollary 5.8.1. *For all $x \in X, x \neq 0$, then there exists $\phi \in X^*$ such that*

$$\|\phi\| = 1 \quad \text{and} \quad \phi(x) = \|x\|$$

Corollary 5.8.2. *For ang $x, y \in X$, if $\phi(x) = \phi(y)$ for all $\phi \in X^*$, then $x = y$.*

Corollary 5.8.3. *Suppose $x_0 \in X$. Then*

$$x_0 = 0 \iff \phi(x_0) = 0 \quad \forall \phi \in X^*$$

Definition 5.9. *Since X^* is a Banach space, we can consider its dual space*

$$(X^*)^* = X^{**}$$

called the <u>bidual</u> of X.

Claim: X may be viewed a subspace of X^{**}.

Proof. For every $x \in X$, we can define an embedding $F : X^* \to X^{**}$

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

That is,

$$F_x(\phi) = \phi(x)$$

Notice that F_x is linear, and

$$|F_x(\phi)| = |\phi(x)| \leq \|\phi\|\,\|x\|$$

$$\implies \|Fx\| \leq \|x\| < \infty \implies F_x \in X^{**}$$

Therefore, the mapping $x \to F_x$ is an embedding from X into X^{**}, and if $x \neq 0$, by the previous corollary, there exists $\phi \in X^*$ such that

$$\|\phi\| = 1 \text{ and } \phi(x) = \|x\|$$

So

$$|F_x(\phi)| = |\phi(x)| = \|x\| = \|x\|\,\|\phi\| \implies \|F_x\| = \|x\|\,\forall x \in X$$

Thus, $X \subset X^{**}$. ∎

Definition 5.10. *If $X = X^{**}$, then we say that X is <u>reflexive</u>*

Example 5.11.

$$(L^p)^{**} = (L^q)^* = L^p \text{ for } 1 < p < \infty$$

5.2 Weak Convergence

Motivation: In finite dimensional normed linear space, every bounded sequence has a convergent subsequence. This property fails in ∞-dimensional spaces.

Example 5.12. *$X = \ell^\infty$ is the space of all bounded sequences. If we define*

$$e_n = (0, 0, \dots, \underbrace{1}_{nth\ term}, 0, 0, \dots$$

where $\|e_n\| = 1$, but $\{e_n\}_{n=1}^\infty$ has no convergent subsequence.

To overcome this difficulty, we introduce another kind of convergence.

Definition 5.13. *Let $(X, \|\cdot\|)$ be a normed linear space. Let $\{x_n\}_{n=1}^\infty$ be a sequence in X. We say*

1. *x_n converges strongly to x if*
$$\|x_n - x\| \to 0$$
denoted $x_n \to x$

2. *x_n converges weakly to x if*
$$\phi(x_n) \to \phi(x)$$
for all $\phi \in X^*$. Denoted $x_n \rightharpoonup x$.

Note 5.14. *When dim$(X) < \infty$, then*

$$\begin{aligned} x_n \rightharpoonup x &\iff \phi(x_n) = \phi(x) \forall \phi \in X^* \\ &\iff w_i(x_n) = w_i(x)\ \forall 1 \le i \le dim(X) \\ &\iff \{x_n\}_{n=1}^\infty \text{ converges to } x \text{ for each coordinate.} \\ &\iff x_n \to x \end{aligned}$$

Proposition 5.15. 1. *If $\{x_n\}_{n=1}^\infty$ is weakly convergent in X, then the weak limit of $\{x_n\}_{n=1}^\infty$ is unique.*

2. *If $x_n \to x$ in X, then $x_n \rightharpoonup x$ in X.*

Proof.

$$|\phi(x_n) - \phi(x)| \le \|\phi\| \, \|x_n - x\| \to 0$$

■

Example 5.16. *When dim$(X) = \infty$, then$x_n \rightharpoonup x \;\not\Longrightarrow\; x_n \to x$. We especially see this when we take $X = L^2([0,1])$, and define the sequence $x_n = \sin(n\pi t)$. Then for all $\phi \in X^* = X$, then*

$$\phi(x_n) = \int_0^1 \phi(t) \sin(n\pi t)\ dt \to 0$$

by the Riemann-Lebesgue Lemma. Therefore, $x_n \to 0$ But

$$\|x_n - 0\| = \|x_n\| = \frac{1}{\sqrt{2}} \not\to 0 \implies x_n \not\to 0$$

Definition 5.17. *Let X be a normed, linear space. The weak topology on X is defined as the weakest topology which makes all of the functionals $\phi \in X^*$ continuous. Let $\Phi := \{\phi^{-1}(U) : U \subset \mathbb{R} \text{ open }, \phi \in X^*\}$. Since*

$$\bigcup \phi^{-1}(U) = X$$

there exists a unique topology $\mathcal{T}_{weak}$ on X having Φ as its sub-base. Moreover,

$$\mathcal{T}_{weak} = \bigcup_{arbitrary} \bigcap_{finite} \phi^{-1}(U)$$

Remark 5.18.

$$x_n \rightharpoonup x \text{ in } X \iff \phi(x_n) \to \phi(x) \text{ for all } \phi \in X^* \iff x_n \to x \text{ in the weak topology } \mathcal{T}_{weak}$$

Lemma 5.19 (Mazur)**.** *Suppose $x_n \rightharpoonup x$ in a Banach space $(X, \|\cdot\|)$. Then, for all $\epsilon > 0, exists \lambda_i \geq 0$ with $1 \leq i \leq n$ with*

$$\sum_{i=1}^{n} \lambda_i = 1$$

such that

$$\left\| x_0 - \sum_{i=1}^{n} \lambda_i x_i \right\| < \epsilon$$

5.3 Convergence in X^*

Definition 5.20. *Let $\{\phi_n\}_{n=1}^{\infty}$ be a sequence in X^* and $\phi \in X^*$.*

1. *Strong Convergence: $\phi_n \to \phi$ strongly in X^* provided $\|\phi_n - \phi\|_{X^*} \to 0$*
2. *Weak Convergence: $\phi_n \rightharpoonup \phi$ weakly in X^* means for all $F \in X^{**}, F(\phi_n) \to F(\phi)$.*

Note 5.21. *Since $X \subset X^{**}$ where every $x \in X$ can be identified with a functional $F_x(\tilde{\phi}) = \tilde{\phi}(x)$. So if $\phi_n \rightharpoonup \phi$ in X^*, then for all $x \in X$*

$$F_x(\phi_n) \to F_x(\phi) \implies \phi_n(x) \to \phi(x) \implies \text{ pointwise}$$

Definition 5.22. *Weak-* Convergence $\phi_n \overset{*}{\rightharpoonup} \phi$ in X if $\phi_n(x) \to \phi(x)$ for all $x \in X$.*

Remark 5.23.

1. *$\phi_n \rightharpoonup \phi$ in $X^* \implies \phi_n \overset{*}{\rightharpoonup} \phi$ in X^*. If $X = X^{**}$, then*

$$\phi_n \rightharpoonup \phi \iff \phi_n \overset{*}{\rightharpoonup} \phi$$

2. *The Weak-* topology on X^* is defined as the weakest topology that makes all the maps $F_x : X^* \to \mathbb{R}$ continuous. That is,*

$$\begin{aligned} X^* &\to \mathbb{R} \\ \phi &\to \phi(x) \end{aligned}$$

is continuous for all $x \in X$.

3. *If $\phi_n \overset{*}{\rightharpoonup} \phi$, then $\{\|\phi_n\|_{X^*}\}$ is bounded and*

$$\|\phi\|_{X^*} \leq \liminf_{n\to\infty} \|\phi_n\|_{X^*}$$

Theorem 5.24 (Banach-Alaoglu)**.** *Let $X^* = B(X, \mathbb{R})$ be the dual space of a normed space X. Then the closed unit ball*

$$B_1(X^*) := \{\phi \in X^* : \|\phi\|_{X^*} \leq 1\}$$

is compact for the weak- topology.*

Proof. (Sketch)
For every $x \in X$, let $D_x := \{y \in \mathbb{R} : |y| \leq \|x\|\} = [-\|x\|, \|x\|]$. Let

$$D = \prod_{x \in X} D_x$$

be the product space. By the Tychonoff Theorem, D is compact with respect to $\mathcal{T}_{\text{product}}$. Now for all $\phi \in B_1(X^*) \implies \|\phi\|_{X^*} \leq 1$. So for all $x \in X$

$$|\phi(x)| \leq \|\phi\|_{X^*} \|x\| \leq \|x\| \implies \phi(x) \in D_x$$

That is, ϕ gives a map $x \to \phi(x) \in D_x \implies \phi \in D$. Therefore, $B_1(X^*) \subset D$.

Now, recall $\phi_n \overset{*}{\rightharpoonup} \phi \iff \phi_n(x) \to \phi(x) \forall x \in X \iff \phi_n \to \phi$ in $(D, \mathcal{T}_{\text{product}})$. Also, if $\phi_n \to \phi$ in D, $\phi_n \in B_1(X^*)$, then $\phi_n(x) \to \phi(x)$ for all x and $|\phi_n(x)| \leq \|x\|$. Therefore $|\phi(x)| \leq \|x\| \implies \phi \in B_1(X^*) \implies B_1(X^*)$ is a closed subset of $D \implies B_1(X^*)$ is compact. ■

Theorem 5.25. *Let X be a separable normed space, then any bounded sequence $\{\phi_n\}_{n=1}^{\infty}$ in X^* has a weak-* convergence subsequence.*

Proof. Since X is separable, it has a countable, dense subset $\{x_m\}_{m=1}^{\infty}$. Since $\{\phi_n\}_{n=1}^{\infty}$ is bounded, then for each m

$$\{\phi_n(x_m)\}_{m=1}^{\infty} \subset \mathbb{R}$$

is a bounded subset of $\mathbb{R}$. Employing a diagonalization argument, there exists a subsequence $\{\phi_{n_k}\}_{k=1}^{\infty} \subset \{\phi_n\}_{n=1}^{\infty}$ is convergent in $\mathbb{R}$ for all m. Since $\{x_m\}_{m=1}^{\infty}$ is dense in X, and $\{\phi_n\}_{n=1}^{\infty}$ is bounded, we want to verify that $\{\phi_{n_k}\}_{k=1}^{\infty}$ is convergent for all $x \in X$. We can do this by showing it's a Cauchy sequence:

$$\begin{aligned}
|\phi_{n_k}(x) - \phi_{n_\ell}(x)| &\leq |\phi_{n_k}(x) - \phi_{n_k}(x_m)| + |\phi_{n_k}(x_m) - \phi_{n_\ell}(x_n)| + |\phi_{n_\ell}(x_m) - \phi_{n_\ell}(x)| \\
&\leq \|\phi_{n_k}\| \|x - x_m\| + |\phi_{n_k}(x_m) - \phi_{n_\ell}(x_n)| + \|\phi_{n_\ell}\| \|x - x_m\|
\end{aligned}$$

Defining $\phi(x) := \lim_{k\to\infty} \phi_{n_k}(x)$, then we have a constructive way to generate the converged candidate. All that's left is to check $\phi \in X^*$ and $\phi_{n_k} \overset{*}{\rightharpoonup} \phi$. ■

6 Hilbert Spaces

Recall that $\mathbb{R}^n$ and $\mathbb{C}^n$ have some nice properties: you have geometrically some measure of distance, as well as algebraically a vector space, and lastly some idea of an angle between vectors. We can extend distance measure via a metric space, and extend the vector space into a normed linear space. But what of angle? Answer: An inner product space! These relationships are illustrated as follows:

$$\begin{pmatrix} \text{distance} & \to & \text{metric space} & \to & \text{topological space} \\ & & \cup & & \cup \\ \text{vector space} & \to & \text{normed linear space} & \to & \text{Banach Space} \\ & & \cup & & \cup \\ \text{angle} & \to & \text{inner product space} & \to & \text{Hilbert Space} \end{pmatrix}$$

6.1 Inner Product Spaces

Definition 6.1. *An* <u>*inner product space*</u> *on a complex linear space X is a map $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{C}$ such that for all $x, y, z \in X$, and $\lambda, \mu \in \mathbb{C}$:*

1. $\langle x, \lambda y + \mu y \rangle = \lambda \langle x, y \rangle + \mu \langle x, y \rangle$
2. $\langle x, y \rangle = \overline{\langle y, x \rangle}$ *(Hermitian symmetric)*
3. $\langle x, x \rangle \geq 0$ *and* $\langle x, x \rangle = 0 \iff x = 0$

A linear space with an inner product is called an <u>*inner product space*</u> *or a* <u>*pre-Hilbert space*</u>.

Remark 6.2. *From (1), and (2),*

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

So it's an anti-linear or conjugate linear in the 1st argument.

Remark 6.3. *In some books, inner product space are required to be linear in the first argument and conjugate linear in the second argument.*

Remark 6.4.

$$\langle x, 0 \rangle = \langle x, 0 \cdot a \rangle = 0 \langle x, a \rangle = 0$$
$$\langle \lambda x, \mu y \rangle = \overline{\lambda} \mu \langle x, 0 \rangle$$

Question 6.5. *What is the relation between inner product spaces and normed linear spaces?*

Proposition 6.6. *Any inner product space $(X, \langle \cdot, \cdot \rangle)$ is a normed linear space $(X, \|\cdot\|)$, where for any $x \in X$:*

$$\|x\| = \sqrt{\langle x, x \rangle}$$

Proof.

- $$\|x\| = \sqrt{\langle x, x \rangle} \geq 0$$
  $$\|x\| = 0 \iff \sqrt{\langle x, x \rangle} = 0 \iff \langle x, x \rangle = 0 \iff x = 0$$

- $$\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{\overline{\lambda} \lambda \langle x, x \rangle} = \sqrt{|\lambda|^2 \langle x, x \rangle} = |\lambda| \sqrt{\langle x, x \rangle}$$

- $$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\
&= \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2 \\
&\leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 && \text{Cauchy Schwarz} \\
&= (\|x\| + \|y\|)^2
\end{aligned}$$

■

Lemma 6.7 (Cauchy-Schwarz)**.** *Let X be an inner product space. Then for any $x, y \in X$*

$$|\langle x,y\rangle| \leq \|x\|\,\|y\|$$

Proof. If is trivial when $y = 0$. So we may assume $y \neq 0$. Then for

$$\lambda = \frac{-\langle x,y\rangle}{\|y\|^2} \in \mathbb{C}$$

we have

$$\begin{aligned}
0 &\leq \langle x+\lambda y, x+\lambda y\rangle \\
&= \|x\|^2 + \overline{\lambda}\langle y,x\rangle + \lambda\langle x,y\rangle + |\lambda|^2\|y\|^2 \\
&= \|x\|^2 - \frac{|\langle y,x\rangle|^2}{\|y\|^2} - \frac{|\langle x,y\rangle|^2}{\|y\|^2} + \frac{|\langle y,x\rangle|^2}{\|y\|^2} \\
&= \|x\|^2 - \frac{|\langle y,x\rangle|^2}{\|y\|^2} \\
&\implies 0 \leq \|x\|^2\|y\|^2 - |\langle y,x\rangle|^2
\end{aligned}$$

■

Corollary 6.7.1. *Let X be an inner product space. Then, the inner product $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{C}$ is continuous.*

Proof. Suppose $x_n \to x, y_n \to y$ in X. Then $\{\|x_n\|\}_{n=1}^{\infty}$, $\{\|y_n\|\}_{n=1}^{\infty}$ are bounded by some constant $M > 0$. Thus:

$$\begin{aligned}
|\langle x_n, y_n\rangle - \langle x,y\rangle| &= |\langle x_n - x, y_n\rangle + \langle x, y_n - y\rangle| \\
&\leq \|x_n - x\|\,\|y_n\| + \|x\|\,\|y_n - y\| \\
&\leq M\|x_n - x\| + \|x\|\,\|y_n - y\| \to 0 \text{ as } n \to \infty
\end{aligned}$$

■

Definition 6.8. *A complete inner product space is called a* <u>Hilbert Space</u>*. So every Hilbert Space is a Banach Space. A common notation for inner products in Hilbert Space $\langle\cdot,\cdot\rangle$ whereas an inner product space only has $(\cdot,\cdot)$.*

Example 6.9. *In $\mathbb{R}^n$, $x = (x_1, \ldots, x_n)$, then*

$$\langle x,y\rangle = x \cdot y$$

Example 6.10. *In $\mathbb{C}^n$, $x = (x_1, \ldots, x_n)$, then*

$$\langle x,y\rangle = \sum_{i=1}^{n} \overline{x}_i y_i$$

Example 6.11. *In $\mathbb{C}^{m\times n}$:= the space of all $m \times n$ matrices with complex entries. Let*

$$A = (a_{i,j})_{1\le j\le n}^{1\le i\le m}, \quad B = (b_{i,j})_{1\le j\le n}^{1\le i\le m} \in \mathbb{C}^{m\times n}$$

Then

$$\langle A, B\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} \overline{a}_{ij} b_{ij} = tr(A^* B)$$

where A^* is the Hermitian conjugate. The corresponding norm is

$$\|A\| = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2}$$

is called the Hilbert-Schmidt norm.

Example 6.12.

$$\ell^2(\mathbb{Z}) := \{\{z_n\}_{n=-\infty}^{\infty} : \sum_{n=-\infty}^{\infty} |z_n|^2 < \infty\}$$

Then

$$\langle x, y\rangle = \sum_{n=-\infty}^{\infty} \overline{x_n} y_n$$

Example 6.13.

$$L^2(\Omega) := \{f : \Omega \to C : \int_\Omega |f|^2 \ dx < \infty\}$$

Then

$$\langle f, g\rangle = \int_\Omega \overline{f(x)} g(x) \ dx$$

defines an inner product on $L^2(\Omega)$. Moreover, $L^2(\Omega)$ is Hilbert.

Question 6.14. *Is $L^p(\Omega), p \neq 2$ an inner product space. Answer: No.*

Theorem 6.15. *A normed linear space X is an inner product space with a normed derived from the inner product if and only if the norm satisfies the parallelogram identity:*

$$\|x + y\| + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2, \ \forall x, y \in X$$

6.2 Orthogonality

Let X be an inner product space. Recall from vector calculus, for any non-zero vectors $x, y \in X$, we can define the angle to be

$$\theta = \cos^{-1}\left(\frac{|(x, y)|}{\|x\| \, \|y\|}\right)$$

In particular, if $(x, y) = 0$, then $\theta = \frac{\pi}{2}$, and we say x and y are orthogonal, and write $x \perp y$.

Definition 6.16. *If $(x, y) = 0$, then x and y are orthogonal, denoted $x \perp y$.*

Proposition 6.17. *Let X be an inner product space. Then the usual theorem of Pythagoras holds. That is, for any $x \perp y$, then*

$$\|x\|^2 + \|y\|^2 = \|x + y\|^2$$

Proof.

$$\|x + y\|^2 = (x + y, x + y) = (x, x) + (x, y) + (y, x) + (y, y) = (x, x) + (y, y) = \|x\|^2 + \|y\|^2$$

■

Definition 6.18. *Suppose $M \subset X$. If $x \perp y$ for all $y \in M$, then we say that x is orthogonal to M, and write $x \perp M$.*

Proposition 6.19. *If $x \perp M$, then $x \perp span(M)$, where*

$$span(M) := \left\{ \sum_{i=1}^{n} \lambda_i x_i : \lambda_i \in \mathbb{C}, x_i \in M \right\}$$

Definition 6.20. *If M is a subset of an inner product space X, then we can define the orthogonal complement of M by*

$$M^{\perp} := \{x \in X : x \perp M\}$$

Note 6.21.

$$M \cap M^{\perp} \subset \{0\}$$

Theorem 6.22. *For any subset M of an inner product space X, $M^{\perp}$ is a closed linear subspace of X.*

Proof.

- *Claim:* M is a linear subspace.
  For any $x_1, x_2 \in M^{\perp}$, then for all $y \in M$, then

  $$(x_1, y) = (x_2, y) = 0$$

  Then for any $\lambda_1, \lambda_2 \in \mathbb{C}$, we see:

  $$(\lambda_1 x_1 + \lambda_2 x_2, y) = 0$$

  Therefore,

  $$\lambda_1 x_1 + \lambda_2 x_2 \in M^{\perp}$$

- *Claim:* M is closed.
  If $\{x_n\}_{n=1}^{\infty} \subset M^{\perp}$ and $x_n \to x$. We want to show $x \in M^{\perp}$. For all $y \in M$, by the continuity of the inner product,

  $$(x, y) = \lim_{n \to \infty} (x_n, y) = 0$$

  Therefore, $x \in M^{\perp}$

■

Theorem 6.23 (Orthogonal Projections)**.** *Let M be a closed linear subspace of a Hilbert space. Then*

1. *For all $x \in H$, there exists a unique element $x_M \in M$ which is closest to x, in the sense that*

   $$\|x - x_M\| = \inf_{y \in M} \|x - y\|$$

2. $(x - x_M) \perp M$

Proof.

- We prove the existence of such a vector that satisifes (1).
  If $x \in M$, then take $x_m = x$. Otherwise, we set $d = \inf_{y \in M} \|x - y\|$ and note that we must have $d > 0$ since $x \notin M$ and M is closed. Consider a minimizing sequence $\{y_n\}_{n=1}^{\infty} \subset M$ such that

  $$\|x - y_n\| \to d \text{ as } n \to \infty$$

  Claim: $\{y_n\}_{n=1}^{\infty}$ is Cauchy.
  Indeed, by the parallelogram law

  $$\|y_n - y_m\|^2 + \|x - y_n + x - y_m\|^2 = 2\|x - y_n\|^2 + 2\|x - y_m\|^2$$

Note, since M is a linear subspace, $\frac{y_n+y_m}{2} \in M$. So

$$\|(x-y_n)+(x-y_m)\| = 2\left\|x - \frac{y_n+y_m}{2}\right\| \geq 2d$$

Therefore,

$$\begin{aligned}\|y_n - y_m\|^2 &= 2\|x-y_n\|^2 + 2\|x-y_m\|^2 - \|(x-y_n)\| + \|(x-y_m)\|^2 \\ &\leq 2\|x-y_n\|^2 + 2\|x-y_m\|^2 - 4d^2 \to 0 \text{ as } m,n\to\infty\end{aligned}$$

Therefore, since $\{y_n\}$ is Cauchy in the Hilbert space, and therefore, $y_n \to x_M \in H$. Since M is closed, $x_M \in M$. Also, since the norm is continuous, we have $\|x - x_M\| = \lim_{n\to\infty}\|x-y_n\| = d$

- Now, we want to show this x_M is unique.
  Suppose another element $\tilde{x}_n \in M$ with
  $$\|x - \tilde{x}_M\| = d$$
  Then
  $$\begin{aligned}\|x_M - \tilde{x}_M\|^2 &= 2\|x-x_M\|^2 + 2\|x-\tilde{x}_M\|^2 - \|2x - x_M + \tilde{x}_M\|^2 \\ &= 2d^2 + 2d^2 - 4\left\|x - \frac{x_M+\tilde{x}_M}{2}\right\|^2 \\ &\leq 2d^2+2d^2-4d^2 = 0\end{aligned}$$
  Therefore, $x_M = \tilde{x}_M$.
- For every $y \in M$, we want to show
  $$(x - x_M, y) = 0$$
  We notice for every $\lambda \in \mathbb{C}$ and $y \in M$ we have
  $$x_M - \lambda y \in M$$
  Thus,
  $$\|x-x_M\|^2 \leq \|x-(x_M-\lambda y)\|^2 = \|x-x_M\|^2 + |\lambda|^2\|y\|^2 + 2Re[\lambda(x-x_M,y)]$$
  So, $|\lambda|^2\|y\|^2 + 2Re[\lambda(x-x_M,y)] \geq 0$. We can make a convenient choice of λ, by let
  $$(x-x_M,y) = re^{i\theta}$$
  Then taking, $\lambda_n = \frac{-1}{n}e^{-i\theta}$, then we get:
  $$\frac{1}{n^2}\|y\|^2 - \frac{2r}{n} \geq 0$$
  $$0 \leftarrow \frac{1}{n}\|y\|^2 \geq 2r \implies r = 0$$

■

Proposition 6.24. *If M is a closed subspace of a Hilbert space H, then*

$$H = M \oplus M^\perp$$

Here, the notation means for every $x \in H$, we can write x uniquely as the decomposition of:

$$x = x_M + x_{M^\perp}, x_M \in M, x_{M^\perp} \in M^\perp$$

We say that H is the direct sum of M and $M^\perp$.

Proof. For every $x \in H$, we choose $x_M \in M$, as the best approximation to x. Then $x_M \in M, x - x_M \in M^\perp$. Therefore,

$$x = x_M + (x - x_M) \in M \oplus M^\perp$$

Therefore,

$$H \subset M \oplus M^\perp \subset H \implies H = M \oplus M^\perp$$

To demonstrate uniqueness, suppose $x = x_M + x_{M^\perp} = \tilde{x}_M + \tilde{x}_{M^\perp}$ Then

$$x_M - x_{M^\perp} = \tilde{x}_M - \tilde{x}_{M^\perp} \in M \cap M^\perp = \{0\}$$

Therefore,

$$x_m = \tilde{x}_M \text{ and } x_{M^\perp} = \tilde{x}_{M^\perp}$$

■

Theorem 6.25 (Properties of the Projection Map)**.** *We proved for every closed linear subspace $M \subset \mathcal{H}$, we ca define the map $P : \mathcal{H} \to \mathcal{H}$ by*

$$x \to x_M \in M$$

such that $x - x_M \perp M$. We consider the properties of this operator P:

1. *$P : \mathcal{H} \to \mathcal{H}$ is linear.*
2. *Range$(P) = M$*
3. *$P^2 = P$*

Proof. 1. For all $x, \tilde{x} \in \mathcal{H}$, we consider two elements:

$$x = x_M + x_M^\perp \in M \oplus M^\perp$$

$$\tilde{x} = \tilde{x}_M + \tilde{x}_M^\perp \in M \oplus M^\perp$$

Then for all $\mu, \lambda \in \mathbb{C}$,

$$P(\lambda x + \mu \tilde{x}) = \lambda x_M + \mu \tilde{x}_M = \lambda P(x) + \mu P(\tilde{x})$$

2. Left as an exercise
3. For all $x \in \mathcal{H}, P^2 x = P(Px) + P(x_M) = P(x)$

■

Definition 6.26. *A projection $P : \mathcal{H} \to \mathcal{H}$ is called orthogonal if for all $x, y \in \mathcal{H}$,*

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

Proposition 6.27. *Let P be an orthogonal projection on $\mathcal{H}$. Then P is bounded and $\|P\| = \{0, 1\}$*

Proof. • P is bounded.
For all $x \in \mathcal{H}$,

– *Case:* $\|Px\| = 0$, then $\|Px\| \leq 0 \cdot \|x\| \leq \|x\|$.
– *Case:* $\|Px\| \neq 0$, then

$$\|Px\| = \frac{\|Px\|^2}{\|Px\|} = \frac{\langle Px, Py \rangle}{\|Px\|} = \frac{\langle x, P^2 x \rangle}{\|Px\|} = \frac{\langle x, Px \rangle}{\|Px\|} \leq \frac{\|x\| \, \|Px\|}{\|Px\|} = \|x\|$$

$\implies \|Px\| \leq \|x\|$ for all $x \in \mathcal{H}$ and $\|P\| \leq 1$.

- Considerations of Range
  If $range(P) = \{0\}$, then $P = 0$ and $\|P\| = 0$. If $range(P) \neq \{0\} \implies \exists x$ such that $Px \neq 0$, then $P(P(x)) = Px$
  $$\implies \|P(Px)\| = \|Px\| \implies \|P\| \geq 1 \implies \|P\| = 1$$

■

Exercise 6.28. *Let P be an orthogonal projection on $\mathcal{H}$. Then*

1. *$I - P$ is an orthogonal projection*
2. *$Ker(P), Range(P)$ are two, closed linear subspaces of $\mathcal{H}$ and $Ker(P) = Range(P)^{\perp}$*

This exercise implies closed linear subspaces and orthogonal projections bare a 1-1 correspondence.

6.3 Orthogonal Bases

Definition 6.29. *Let X be an inner product space. A subset $S = \{e_\alpha : \alpha \in A\} \subset X$ is called orthonormal if it satisfies*

1. *(Orthogonality) $e_\alpha \perp e_\beta$ for all $\alpha, \beta \in A, \alpha \neq \beta$.*
2. *(Normalized) $\|e_\alpha\| = 1$ for all $\alpha \in A$*

In other words, $\langle e_\alpha, e_\beta \rangle = \delta_{\alpha,\beta} = \begin{cases} 1 & \alpha = \beta \\ 0 & \alpha \neq \beta \end{cases}$

We can also go about a different approach than the book to define orthonormal bases:

Theorem 6.30 (Bessel's Inequality)**.** *Let $S = \{e_\alpha : \alpha in A\}$ be an orthonormal set of an inner product space X. Then for all $x \in X$, we have*

1. *There are at most countably many $\alpha \in A$ with $\langle e_\alpha, x \rangle \neq 0$*
2. *(Bessel's Inequality)*
   $$\sum_{\alpha \in A} |\langle e_\alpha, x \rangle|^2 \leq \|x\|^2$$

Proof. We first prove the inequality for any finite subset $A_f \subset A$. Without loss of generality, we may enumerate A_f as
$$A_f := \{1, 2, \dots, n\}$$
and we need to prove $\sum_{i=1}^n |\langle e_i, x \rangle|^2 \leq \|x\|^2$ Observe:
$$\begin{aligned}
0 &\leq \left\| x - \sum_{i=1}^n \langle e_i, x \rangle e_i \right\|^2 \\
&= \langle x - \sum_{i=1}^n \langle e_i, x \rangle e_i, x - \sum_{j=1}^m \langle e_j, x \rangle e_j \rangle \\
&= \langle x, x \rangle - \sum_{i=1}^n \overline{\langle e_i, x \rangle} \langle e_i, x \rangle - \sum_{j=1}^m \langle e_j, x \rangle \langle x, e_j \rangle + \sum_{i=1}^n \sum_{j=1}^m \overline{\langle e_i, x \rangle} \langle e_j, x \rangle \langle e_i, e_j \rangle \\
&= \|x\|^2 - \sum_{i=1}^n |\langle e_i, x \rangle|^2
\end{aligned}$$

So for all finite subset $A_f \subset A$

$$\sum_{i=1}^{n} |\langle e_i, x\rangle|^2 \leq \|x\|^2$$

Now for any $n \in \mathbb{N}$, let $A_n := \{\alpha \in A : |\langle e_\alpha, x\rangle| \geq \frac{1}{n}\}$. Then the possible number of combinations A_n can be bounded

$$\# A_n \leq n^2 \|x\|^2$$

since otherwise, we can pick A_f too be $[n^2 \|x\|^2 + 1]$ many elements in A_n then

$$\frac{[n^2 \|x\| + 1]}{n^2} \leq \sum_{\alpha \in A_f} |\langle e_\alpha, x\rangle|^2 \leq \|x\|^2$$

then the set $\{\alpha \in A : \langle e_\alpha, x\rangle \neq 0\} = \bigcup_{n=1}^{\infty} A_n$ is countable. Therefore, for all n

$$\sum_{i=1}^{n} |\langle e_i, x\rangle|^2 \leq \|x\|^2$$

which gives

$$\sum_{\alpha \in A} |\langle e_\alpha, x\rangle|^2 = \sum_{i=1}^{\infty} |\langle e_i, x\rangle|^2 \leq \|x\|^2$$

■

Corollary 6.30.1. *Let $U := \{e_\alpha : \alpha \in A\}$ be an orthogonal subset of a Hilbert space $\mathcal{H}$. Then*

1. $\sum_{\alpha \in A} \langle e_\alpha, x\rangle e_\alpha \in \mathcal{H}$
2. $x - \sum_{\alpha \in A} \langle e_\alpha e_\alpha \in U^\perp$
3. $\|x\|^2 = \left\|x - \sum_{\alpha \in A} \langle e_\alpha e_\alpha\right\|^2 + \left\|\sum_{\alpha \in A} \langle e_\alpha e_\alpha\right\|^2$

Proof. 1. For all $x \in \mathcal{H}$, there are at most countably many $\alpha \in A$ with

$$\langle e_\alpha, x\rangle \neq 0$$

Without loss of generality, we may assume $\alpha \in \mathbb{N}$. So

$$\sum_{\alpha \in A} \langle e_\alpha, x\rangle e_\alpha = \sum_{i=1}^{n} \langle e_i, x\rangle e_i$$

We want to show that the sequence of partial sums

$$\left\{ x_m = \sum_{i=1}^{m} \langle e_i, x\rangle e_i \right\} \text{ converges in } \mathcal{H}$$

By Bessel's inequality, we know

$$\sum_{i=1}^{m} |\langle e_i, x\rangle|^2$$

is convergent. So whenever $m > n$

$$\|x_m - x_n\|^2 = \left\| \sum_{i=n+1}^{m} \langle e_i, x\rangle e_i \right\|^2 = \sum_{i=n+1}^{m} |\langle e_i, x\rangle|^2 \to 0 \text{ as } n \to \infty$$

Therefore, $\{x_m\}_{m=1}^{\infty}$ is Cauchy in $\mathcal{H}$ and thus is convergent.

$$\sum_{\alpha \in A} \langle e_\alpha, x\rangle e_\alpha = \lim_{m \to \infty} x_m \in \mathcal{H}$$

2. Notice, for all $\beta \in V$,

$$\left\langle \sum_{\alpha \in A} \langle e_\alpha, x \rangle e_\alpha, e_\beta \right\rangle = \lim_{m \to \infty} \left\langle \sum_{i=1}^{m} \langle e_i, x \rangle e_i, e_\beta \right\rangle = \lim_{m \to \infty} \sum_{i=1}^{m} \overline{\langle e_i, x \rangle} \delta_{i,\beta} = \sum_{i=1}^{\infty} \overline{\langle e_i, x \rangle} \delta_{i,\beta} = \overline{\langle e_\beta, x \rangle}$$

Therefore,

$$\left\langle x - \sum_{\alpha \in A} \langle e_\alpha, x \rangle e_\alpha, e_\beta \right\rangle = \langle x, e_\beta \rangle - \overline{\langle e_\beta, x \rangle} = 0$$

3. More over, since

$$\|x - x_m\|^2 = \|x\|^2 - \sum_{i=1}^{m} |\langle e_i, x \rangle|^2$$

Let $m \to \infty$, then

$$\left\| x - \sum_{\alpha \in A} \langle e_\alpha, x \rangle e_\alpha \right\|^2 = \|x\|^2 - \left\| \sum_{\alpha \in A} \langle e_\alpha, x \rangle e_\alpha \right\|^2$$

■

Question 6.31. *When does the equality in Bessel's Inequality hold?*

Theorem 6.32. *Suppose $U = \{e_\alpha : \alpha \in A\}$ is an orthogonal set in a Hilbert space $\mathcal{H}$. Then the following statements are equivalent:*

1. *U is an orthonormal basis of $\mathcal{H}$ in the sense that for all $x \in \mathcal{H}$,*

$$x = \sum_{\alpha \in A} \langle e_\alpha, x \rangle e_\alpha$$

2. *U is complete in the sense that $U^\perp$, or equivalently, U is a maximal orthogonal set or $\langle e_\alpha, x \rangle = 0$ for all $\alpha \in A \implies x = 0$.*

3. *Parseval's Identity Holds: $\|x\|^2 = \sum_{\alpha \in A} |\langle e_\alpha, x \rangle|^2$*

Proof. (1) ⇒ (2) For all $x \in U^\perp$, then by (1), we know

$$x = \sum_{\alpha \in A} \langle e_\alpha, x \rangle e_\alpha = \sum_{\alpha \in A} 0 e_\alpha = 0$$

(2) ⇒ (3) For all $x \in \mathcal{H}$, then

$$\|x\|^2 - \sum_{\alpha \in A} |\langle e_\alpha, x \rangle|^2 = \left\| x - \sum_{\alpha \in A} \langle e_\alpha, x \rangle \right\|^2 = \|0\|^2 = 0$$

(3) ⇒ (1) For all $x \in \mathcal{H}$, then

$$\left\| x - \sum_{\alpha \in A} \langle e_\alpha, x \rangle e_\alpha \right\|^2 = \|x\|^2 - \sum_{\alpha \in A} |\langle e_\alpha, x \rangle|^2 = 0$$

■

Theorem 6.33 (Generalized Parseval's Identity). *Suppose that $U = \{e_\alpha : \alpha \in A\}$ is an orthonormal basis of a Hilbert space $\mathcal{H}$. For any $x, y \in \mathcal{H}$*

$$x = \sum_{\alpha \in A} a_\alpha e_\alpha \text{ for } a_\alpha = \langle e_\alpha, x \rangle$$

$$y = \sum_{\alpha \in A} b_\alpha e_\alpha \text{ for } b_\alpha = \langle e_\alpha, x \rangle$$

Then

$$\langle x, y \rangle = \sum_{\alpha \in A} \overline{a_\alpha} b_\alpha$$

Definition 6.34. *Let $(X_1, (\cdot,\cdot)_1)$ and $(X_2, (\cdot,\cdot)_2)$ be two inner product spaces. If there exists an isomorphism $T : X_1 \to X_2$ such that for all $x, y \in X$:*

$$(Tx, Ty)_2 = (x, y)_1$$

Then we say that the inner product space X_1, X_2 are isomorphic.

Remark 6.35. *Parseval's Identity says that a Hilbert space $\mathcal{H}$ with an orthonormal basis $\{e_\alpha : \alpha \in A\}$ is isometric to the sequence space $\ell^2(A)$, with an isomorphism defined*

$$\mathcal{H} \to \ell^2(A)$$

$$x \to \langle e_\alpha, x \rangle$$

In particular, if A is finite with $\#A = N$, then $\mathcal{H}$ is isomorphic to $\mathbb{C}^N = \ell^2([1, 2, \dots, N])$. If A is countable (i.e. $\mathcal{H}$ is separable, ∞-dimensional), then $\mathcal{H}$ is isomorphic to $\ell^2(\mathbb{N}) = \ell^2$.

Proof. (of Parseval's Identity)
For all $x, y \in \mathcal{H}$, there are at most countably many $\alpha \in A$ such that

$$\langle e_\alpha, x \rangle \neq 0 \text{ or } \langle e_\alpha, y \rangle \neq 0$$

Without loss of generality, we may assume

$$x = \sum_{i=1}^{\infty} a_i e_i, \quad y = \sum_{i=1}^{\infty} b_i e_i$$

Then by the continuity of $\langle \cdot, \cdot \rangle$,

$$\langle x, y \rangle = \lim_{m \to \infty} \langle \sum_{i=1}^{m} a_i e_i, \sum_{i=1}^{m} b_i e_i \rangle = \lim_{m \to} \bar{a}_i b_i = \sum_{\alpha \in A} \bar{a}_\alpha b_\alpha$$

To justify the last equality, we need convergence. This can be seen by

$$\sum_{i=1}^{\infty} |\overline{a_i} b_i| \leq \sqrt{\sum_{i=1}^{\infty} |\overline{a_i}|^2} \sqrt{\sum_{i=1}^{\infty} |b_i|^2} = \|x\|^2 \|y\|^2 < \infty$$

∎

Question 6.36. *Does every Hilbert Space have an orthonormal basis?*

Definition 6.37. *$U := \{e_\alpha : \alpha \in A\}$ is an orthonormal basis of $\mathcal{H}$ if every $x \in \mathcal{H}$ can be uniquely represented by*

$$x = \sum_{\alpha \in A} \langle e_\alpha, x\rangle e_\alpha$$

and $\langle e_i, , e_j\rangle = \delta_{i,j}$.

Example 6.38. *$\mathbb{C}^n$ equipped with the inner product $\langle x, y\rangle = \sum_{i=1}^n \overline{x}_i y_i$. We define:*

$$e_i = (0, 0, \dots, \underbrace{1}_{ith}, \dots, 0)$$

Then

$$\langle e_i, e_j\rangle = \delta_{i,j} \implies \text{ orthonormal}$$

And for all $x = (x_1, \dots, x_n) \in \mathbb{C}^n$, then

$$x = \sum_{i=1}^n x_i e_i = \sum_{i=1}^n \langle e_i, x\rangle e_i$$

Therefore, $\{e_i\}_{i=1}^n$ is an orthonormal basis of $\mathbb{C}^n$.

Remark 6.39. *$\{\langle e_\alpha, x\rangle : \alpha \in A\}$ is called the Fourier Coefficients of x with respect to the basis $\{e_\alpha\}$*

Example 6.40. *Consider $L^2([0, 2\pi])$ equipped with $\langle f, g\rangle = \int_0^{2\pi} \overline{f(x)} g(x)\ dx$. Define*

$$e_n(x) := \frac{1}{\sqrt{2\pi}} e^{inx} \text{ for } n \in \mathbb{Z}$$

To show this is orthonormal,

$$\begin{aligned}
\langle e_n(x), e_m(x)\rangle := \frac{1}{2\pi}\int_0^{2\pi} e^{-inx} e^{imx} dx \\
&= \frac{1}{2\pi}\int_0^{2\pi} e^{i(m-n)x}\ dx \\
&= \begin{cases} 1 & m = n \\ \frac{e^{i(m-n)x}}{i(m-n)} = 0 & m \neq n \end{cases} \\
&= \delta_{m,n}
\end{aligned}$$

In Chapter 7, we will show that this particular $\{e_n(x) : n \in \mathbb{Z}\}$ is complete. Therefore, we will continue believing $\{e_n(x) : n \in \mathbb{Z}\}$ is an orthonormal basis. Therefore, for any $f \in L^2([0, 2\pi])$

$$\langle e_n, f\rangle = \frac{1}{\sqrt{2\pi}}\int_0^{2\pi} f(x) e^{-inx}\ dx = \hat{f}_n$$

Example 6.41. *Let $D := \{z \in \mathbb{C} : |z| < 1\}$ be the open unit disk in $\mathbb{C}$. Let*

$$H^2(D) := \{u : D \to \mathbb{C} : u \text{ is analytic on } D \text{ and } \iint_D |u(x+iy)|^2 dxdy < +\infty\}$$

and equip this space with the inner product

$$\langle u, v\rangle := \iint_D \overline{u(z)} v(z)\ dxdy$$

Let $e_n(z) = \sqrt{\frac{n}{\pi}} z^{n-1}, n \in \mathbb{N}$ Then $\{e_n(z) : n \in \mathbb{N}\}$ forms an orthonormal basis of $H^2(D)$.

Question 6.42. *So now how do we attempt to prove the existence of an orthonormal basis for an Hilbert space.* $\implies$ *Zorn's Lemma, which is a consequence of the Axiom of Choice.*

Axiom of Choice: Let X be a non-empty set and let $P_0(X)$ be the family of non-empty subset of X. Then there exists a function

$$f : P_0(X) \to \mathbb{X}$$

such that

$$f(A) \in A$$

for any set $A \in P_0(X)$.

Definition 6.43. *A partial order on a set X is a relation "$\preceq$" with the properties*

- *For all $x \in X, x \preceq x$*
- *For all $x, y \in X$, if $x \preceq y$ and $y \preceq x$, then $x = y$*
- *For all $x, y, z \in X$, if $x \preceq y$ and $y \preceq z$, then $x \preceq z$.*

Definition 6.44. *Let $(X, \preceq)$ be a partially ordered set. Then $Y \subset X$ is totally ordered if for every $x, y \in Y$, we have $x \preceq y$ or $y \preceq x$.*

Definition 6.45. *Given $Y \subset (X, \preceq)$ a partially ordered set, we say $y_0 \in X$ is an upper bound for Y if for all $y \in Y, y \preceq y_0$.*

Definition 6.46. *Given $Y \subset (X, \preceq)$ a partially ordered set, we say y_0 is a maximal element for Y if for all $y \in Y, y_0 \preceq y$ implies $y = y_0$.*

Lemma 6.47 (Zorn's Lemma). *Let S be a non-empty, partially ordered set with the property that every totally ordered subset has an upper bound in S. Then S contains a maximal element.*

Theorem 6.48. *Every Hilbert space $\mathcal{H}$ has an orthonormal basis.*

Proof. It is trivial for $\mathcal{H} = \{0\}$. We may assume $\mathcal{H} \neq \{0\}$. Let

$$S = \{U : U = \{e_\alpha\} \text{ is an orthonormal subset of } \mathcal{H}\}$$

If we want to use Zorn's Lemma, we need a partial order on S. Since S is a set of sets, it's probably wise to use inclusion as a partial order. So we define a partial order "$\preceq$" on S by

$$U \preceq V \iff U \subset V$$

Notice, for any totally order subset $\{U_\alpha : \alpha \in A\}$ of S, it must have an upper bound, defined:

$$U := \bigcup_{\alpha \in A} U_\alpha \subset S$$

So by Zorn's Lemma, S has a maximal element U. Now we want to show this U also an orthonormal basis. Clearly it's orthonormal. But we need completeness.

Claim: $U^\perp = \{0\}$.
Suppose otherwise. Then there exists $e \in U^\perp$ with $\|e\| = 1$. Then

$$\tilde{U} := U \cup \{e\}$$

is an orthonormal set with $U \subset \tilde{U}$. But this contradicts the fact that U is maximal.

Therefore, since $U^\perp = \{0\}$, we know U is an orthonormal basis of $\mathcal{H}$. ∎

6.4 How to Construct an Orthonormal Basis

If $\mathcal{H}$ is separable, then we may use the Gram-Schmidt orthonormalization procedure:

Gram-Schmidt
Input: $\{x_1, x_2, \ldots\}$ linearly independent in $\mathcal{H}$
Output: An orthonormal set $\{e_1, e_2, \ldots\}$ such that for all n

$$span\{x_1, x_2, \ldots, x_n\} = span\{e_1, e_2, \ldots, e_n\}$$

Let $y_1 = x_1$, and set $e_1 = \frac{y_1}{\|y_1\|}$.
Let $y_2 := x_2 - \langle e_1, x_2\rangle e_1$, and set $e_2 = \frac{y_2}{\|y_2\|}$
$\vdots$
$y_n := x_n - \sum_{i=1}^{n-1}\langle e_i, x_n\rangle e_i$ and set $e_n = \frac{y_n}{\|y_n\|}$
After nth iteration: Then $\{e_i\}_{i=1}^n$ is orthonormal and

$$span\{x_1, x_2, \ldots, x_n\} = span\{e_1, e_2, \ldots, e_n\}$$

Theorem 6.49. *A Hilbert space $\mathcal{H}$ is separable if and only if it has a countably, orthogonal basis S*

Proof. Without loss of generality, we may assume $H \neq \{0\}$.
($\Leftarrow$) Suppose $\{e_n\}_{n=1}^N, (N \leq \infty)$ is an orthonormal basis of $\mathcal{H}$. Then the set

$$\{x = \sum_{n=1}^{N} a_n e_n : \text{ both } Re(a_n) \text{ and } Im(a_n) \text{ are rationals}\}$$

is a dense subset of $\mathcal{H}$. Therefore, $\mathcal{H}$ is separable.
($\Rightarrow$) Suppose $\{x_n\}_{n=1}^\infty$ is a countable, dense subset of $\mathcal{H}$. Then it must have a linearly independent subset

$$\{y_n\}_{n=1}^N \subset \{x_n\}_{n=1}^\infty$$

such that $N \leq \infty$ and

$$span\{y_n\}_{n=1}^N = span\{x_n\}_{n=1}^\infty$$

Apply Gram-Schmidt to $span\{y_n\}_{n=1}^N$, to construct an orthonormal set

$$\implies \{e_n\}_{n=1}^N \text{ with } span\{e_n\}_{n=1}^N = span\{y_n\}_{n=1}^N = span\{x_n\}_{n=1}^\infty$$

Claim: $\{e_n\}_{n=1}^N$ is an orthonormal basis
Let $U = \{e_n\}_{n=1}^N$. Then

$$\mathcal{H} = \overline{span\{x_n\}_{n=1}^\infty} = \overline{span(U)} = (U^\perp)^\perp$$

$\implies U^\perp = \{0\}$ and U is an orthonormal basis of $\mathcal{H}$. ■

Example 6.50. *Let $\mathcal{H} = L^2([a,b])$ with $\langle f, g\rangle = \int_a^b \overline{f} g\ dx$.*

- *$C([a,b])$ is dense in $L^2(a,b)$*
- *Polynomials are dense in $C([a,b])$*
- *$span\{1, x, x_2, \ldots\} =$ Polynomials*

BUT!! These are not orthonormal to one another!! We need to apply Gram-Schmidt to $\{1, x, x_2, \ldots\}$ to get an orthonormal basis of $\mathcal{H}$ consisting of polynomial functions.

Example 6.51. *On $L^2[-1,1]$, we get the Legendre Polynomials:*

$$P_0(x) = 1$$
$$P_1(x) = x$$
$$P_2(x) = \frac{1}{2}(3x^2 - 1)$$
$$P_3(x) = \frac{1}{2}(5x^3 - 3x)$$

with coefficients:

$$\int_0^1 P_n(x)P_m(x)\ dx = \frac{2}{2n+1}\delta_{m,n}$$

In general, we can consider a measure μ on $[a,b]$, we can define the space:

$$L^2_\mu([a,b]) = \{f : [a,b] \to \mathbb{C} : \int_a^b |f(x)|^2 d\mu < \infty\}$$

with inner product

$$\langle f, g\rangle_\mu = \int_a^b \overline{f} g\ d\mu$$

In particular, for $\mu = w(x)\ dx$ with $w : (a,b) \to (0,\infty)$ continuous then

$$\langle f, g\rangle_w = \int_a^b \overline{f(x)} g(x) w(x)\ dx$$

We can apply Gram-Schmidt, to $\{1, x, x_2, \ldots\}$, we get orthonormal basis of polynomials for $L^2_w([a,b])$

Example 6.52. *On $L^2([-1,1])$*

- $w = \sqrt{1-x^2}$ *results in the Tchebyschev Polynomials*
- $w = e^{\frac{-x^2}{2}}$ *results in the Hermite Polynomials*

Theorem 6.53 (Riemann-Lebesgue Lemma)**.** *Let $f : [0, 2\pi] \to \mathbb{R}$ be continuous. Then*

$$\int_0^{2\pi} f(x)\sin(nx)\ dx \to 0 \text{ as } n \to \infty$$

Before we prove this theorem, we prove a useful lemma:

Lemma 6.54. *Let $\{e_n\}_{n=1}^\infty$ be an orthonormal set in an inner product space X. Then for every $x \in X$,*

$$(e_n, x) \to 0 \text{ as } n \to \infty$$

Proof. By Bessel's Inequality,

$$\sum_{n=1}^\infty |(e_n, x)|^2 \le \|x\|^2 < \infty$$
$$\implies \sum_{n=1}^\infty |(e_n, x)|^2 \text{ convergent}$$
$$\implies |(e_n, x)|^2 \to 0 \text{ as } n \to 0$$
$$\implies (e_n, x) \to 0 \text{ as } n \to 0$$

∎

Proof. (Of Riemann-Lebesgue)
Let $X = L^2([0, 2\pi])$ and define an orthonormal basis

$$\{e_n := \frac{1}{\sqrt{2\pi}} e^{inx} : n \in \mathbb{Z}\}$$

Then for all $f \in L^2([0, 2\pi])$,

$$\langle e_n, f\rangle = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(x)e^{-inx}\ dx \to 0$$

by the previous lemma. Suppose f is real-valued, then:

$$\int_0^{2\pi} f(x)\cos(nx)\ dx \to 0 \text{ and } \int_0^{2\pi} f(x)\sin(nx)\ dx \to 0$$

■

7 Measure Spaces and Measurable Functions

7.1 Motivation

Question 7.1. *Why do we care about measures? Answer: Lots of reasons:*

1. Measures have nice properties - signed measures, product measures, limits of measures, abstract upon integration and differentiation, etc. (Can be roughly viewed as a generalization of functions)
2. Measures lead to new integration theorems, which overcome difficulties of Riemann integration.
3. Probability theory
4. Geometric objects can be viewed as measures, and we can take limits of them, allowing us to studying properties of the limit object.

Motivation: Given a set $A \subset \mathbb{R}^n$, we want to *measure the size* of A.

Example 7.2. *Define*

$$\mu(A) = \begin{cases} Card(A) & A \text{ is finite} \\ \infty & \text{otherwise} \end{cases}$$

Question 7.3. *How about the length / area / volume? Maybe we can use well-defined areas of cubes, balls, etc.*

Question 7.4. *Can we define "volume" $\mu(A)$ for every subset A of $\mathbb{R}^n$.*

Note 7.5. *A meaningful notion $\mu(A)$ of volume shall satisfy some obviously desirable conditions:*

1. $\mu(\emptyset) = 0$
2. *If $A \subset B$, then $\mu(A) \leq \mu(A)$*
3. *(Translation Invariance) $\mu(A + x) = \mu(A)$ for all $x \in \mathbb{R}^n$*
4. $\mu(\text{ cubes}) = volume(\text{ cubes})$
5. $\mu(\bigcup_{i=1}^\infty A_i) = \sum_{i=1} \mu(A_i)$ *where the union is disjoint.*

Returning to the previous question, it is not possible to define $\mu(A)$ for every subset A in a way consistent with the above properties.

Example 7.6 (Vitali). *Assume such a measure μ exists on $\mathbb{R}$. Define an equivalence relation on $\mathbb{R}$ as follows*

$$x \sim y \iff x - y \in \mathbb{Q}$$

Let us pick a single element $x \in [0,1]$ from each equivalence class. All these picked elements x form a set $A \subset [0,1]$.
We want to calculate $\mu(A)$.

- *We know for any $h \in \mathbb{Q}$, we define:*
$$A_h = A + h$$
and by Property 3,
$$\mu(A_h) = \mu(A)$$
- *But also notice that for $h, s \in \mathbb{Q}$, then*
$$A_h \cap A_s = \emptyset$$
Therefore, $\{A_h\}_{h \in \mathbb{Q}}$ is a disjoint sequence of sets. Therefore, applying property 5, we see:
$$\infty = \mu(\mathbb{R}) = \mu(\bigcup_{h \in \mathbb{Q}} A_h) = \sum_{h \in \mathbb{Q}} \mu(A_h) = \sum_{h \in \mathbb{Q}} \mu(A)$$
$$\implies \mu(A) > 0$$
- *On the other hand,*
$$\bigcup_{h \in \mathbb{Q} \cap [0,1]} A_h \subset [0,2]$$
So
$$2 = \mu([0,2]) \geq \sum_{h \in \mathbb{Q} \cap [0,1]} \mu(A_h) = \sum_{h \in \mathbb{Q} \cap [0,1]} \mu(A)$$
$$\implies \mu(A) = 0$$

This is a contradiction by the previous point. Therefore, there is no way to define length consistently for all subset of $\mathbb{R}$.

Idea: Define "length" μ for a "nice" family of subsets and not just all subsets. This leads us to a σ-algebra.

7.2 σ-algebras and Measures

Definition 7.7. *A σ-algebra on a set X is a collection $\mathscr{A}$ of subets of X such that*

1. $\emptyset \in \Sigma$
2. *If $A \in \Sigma$, then $A^c \in \Sigma$*
3. *If $A_1, A_2, \ldots,$ is a countable family of sets in Σ, then*
$$\bigcup_{i=1}^{\infty} A_i \in \Sigma$$

Definition 7.8. *Σ is called an algebra if the third condition is replaced by*

$$\bigcup_{i=1}^{N} A_i \in \Sigma$$

for each N.

Remark 7.9. *$X \in \Sigma$ because $X = \emptyset^c$*

Remark 7.10. *If $A_i \in \Sigma$ for all i, then by de Morgan's Law,*

$$\bigcup_{i=1}^{\infty} A_i = \left(\bigcap_{i=1}^{\infty} A_i\right)^c \in \Sigma$$

Remark 7.11. *If $A, B \in \Sigma$ then*

$$A \setminus B = A \cap B^c \in \Sigma$$

So σ-algebras are "closed" under complementation, countable unions, and countable intersections. Similarly, algebra's are closed under complementation, but are limited to finite unions and intersections.

Example 7.12. *The smallest σ-algebra on X is $\{\emptyset, X\}$. The largest σ-algebra on X is 2^X.*

Example 7.13. *Let $\mathcal{F} \subset 2^X$ be an arbitrary collection of subset of X. Then*

$$\Sigma(\mathcal{F}) = \bigcap \{\Sigma : \mathcal{F} \in \Sigma, \Sigma \text{ a } \sigma\text{-algebra on } X\}$$

is also a σ-algebra. It is the smallest σ-algebra on X that contains $\mathcal{F}$. Is is called the σ-algebra generated by $\mathcal{F}$.

Example 7.14. *Let X be a metric space, or even a topological space. Let $\mathcal{T} \subset 2^X$ be the collection of all open set in X. Then the σ-algebra $\mathscr{B}(X)$ is called the <u>Borel σ-algebra</u> of X. Each element of $\mathscr{B}(X)$ is called a <u>Borel set</u>.*

Elements of $\mathscr{B}(X)$ include:

- All open sets
- All closed sets
- $F_\sigma := \bigcup_{n=1}^{\infty} F_n$, where F_n is closed.
- $G_\delta := \bigcap_{n=1}^{\infty} G_n$, where G_n is open.

Example 7.15. *$[a, b) = [a, a+1] \cap (a-1, b)$ is also Borel.*

Definition 7.16. *A <u>measureable space</u> $(X, \mathscr{A})$ is a set X and a σ-algebra Σ on X. Elements of Σ are called <u>measurable sets</u>.*

Definition 7.17. *A <u>measure</u> μ on X is a function*

$$\mu : \Sigma \to [0, \infty]$$

such that

1. *$\mu(\emptyset) = 0$*
2. *μ is countably additive:*

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

for each countable family of mutually disjoint sets $\{E_n\}_{n=1}^{\infty} \subset \Sigma$.

Remark 7.18. *Countable-additivity is an essential requirement. Without it, one cannot develop a satisfactory theory of integration.*

Lemma 7.19. ***Some Consequences of Countable Additivity***

1. *Finite Additivity:*
$$\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i)$$

2. *If* $A \subset B \implies \mu(A) \leq \mu(B)$

   Proof. Let $B = A \cup (B \setminus A)$. Then
$$\mu(B) = \underbrace{\mu(A) + \mu(B \setminus A)}_{\text{disjoint sets}} \geq \mu(A)$$
∎

3. *Continuity of increasing sequences: If* $A_1 \subset A_2 \subset \ldots \subset A_n \in \mathscr{A}$, *then*
$$\lim_{n\to\infty} \mu(A_n) = \mu\left(\bigcup_{i=1}^{\infty} A_i\right)$$

4. *Continuity of decreasing sequences: If* $A_1 \supset A_2 \supset \ldots \supset A_n \in \mathscr{A}$, *and* $\mu(A_1) < +\infty$, *thn*
$$\lim_{n\to\infty} \mu(A_n) = \mu\left(\bigcap_{i=1}^{\infty} A_i\right)$$

5. *Countable Subadditivity: For any* $\{A_i\} \subset \mathscr{A}$, *then*
$$\mu(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$$

   Proof. Let $B_1 := A_1$, and recursively define:
$$B_n = A_n - \left(\bigcup_{i=1}^{n-1} A_i\right)$$
   Then
$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$$
   which is a disjoint union and $B_i \subset A_i$ for every i. So
$$\mu(\bigcup_{i=1}^{\infty} A_i) = \mu(\bigcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} \mu(B_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$$
∎

Definition 7.20. *A measure space* $(X, \mathscr{A}, \mu)$ *is a set* X, *a* σ*-algebra* $\mathscr{A}$ *on* X, *and a measure* $\mu : \mathscr{A} \to [0, \infty]$.

Definition 7.21. *A measure* μ *is said to be*

- *finite if* $\mu(X) < \infty$
- *σ-finite if there exists a sequence* $\{A_i\} \subset \mathscr{A}$ *such that* $X = \bigcup_{i=1}^{\infty} A_i$ *and* $\mu(A_i) < \infty$ *for all* n.
- *probability measure if* $\mu(X) = 1$.

Lemma 7.22. *If $\mathscr{A}_1 \subset \mathscr{A}_2$ are σ-algebras, and μ is a measure on $\mathscr{A}_2$, then μ is also a measure on $\mathscr{A}_1$.*

Remark 7.23. *It is also useful to consider*

- *signed measures:* $\mu : \mathscr{A} \to [-\infty, \infty]$
- *comples measures:* $\mu : \mathscr{A} \to \mathbb{C}$
- *vector-valued measures:* $\mu : \mathscr{A} \to$ *linear space.*

Example 7.24 (Counting Measure)**.** *The space $(X, 2^X, \nu)$, define*

$$\nu(A) = \begin{cases} Card(A) & A \text{ is finite} \\ \infty & \text{otherwise} \end{cases}$$

Example 7.25 (Dirac Measure)**.** *The space $(X, 2^X, \delta_x)$, where x is a given point in X. Then define*

$$\delta_x(A) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

Example 7.26 (Discrete Measure / Atomic Measure)**.** *The space $(X, 2^X, \mu)$, where $\{x_i\}$ is a set of points in X and $\{m_i\} \subset \mathbb{R}$ are associated weights. Then define*

$$\mu(A) = \sum_{i=1}^{\infty} m_i \delta_{x_i}(A)$$

Definition 7.27. *Given $X, \mathscr{A} = \mathscr{B}(X)$ then Borel σ-algebra. Then any measure μ on $(X, \mathscr{B}(X))$ is called a Borel measure.*

Example 7.28 (Lebesgue Measure)**.** *Given $X = \mathbb{R}^n, \mathscr{A} = \mathscr{B}(X)$. Then for any cube $C = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n]$ define*

$$\lambda(C) = vol(C) = \prod_{i=1}^{n} (b_i - a_i)$$

Then for any Borel set E, then

$$\lambda(E) = \inf \sum_{i=1}^{\infty} \lambda(C_i)$$

where the infimum is taken over all countable coverings by cubes

$$E \subset \bigcup_{j=1}^{\infty} C_j$$

by closed cubes.

Lemma 7.29. ***Properties of Lebesgue Measure:***

1. *Translation Invariance:* $\lambda(x + A) = \lambda(A)$
2. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation, then*

$$\lambda(TA) = |det(T)|\lambda(A)$$

In particular, λ is rotational invariant, and λ has a scaling property.

$$\lambda(tA) = t^n \lambda(A)$$

for any $t \geq 0$.

3. It's σ-finite since $\mathbb{R}^n$ is the countable union of cubes of unit edge length.

Definition 7.30. *Let $(X, \mathscr{A}, \mu)$ be a measure space and $A \subset X$ is said to have <u>measure zero</u> if it is measurable and $\mu(A) = 0$.*

Example 7.31.

$$\lambda(\{x_0\}) = 0$$
$$\lambda(\mathbb{Q}) = 0$$
$$\lambda(\text{ Cantor Set }) = 0$$

Remark 7.32. *Suppose A is a measure zero set in a measure space $(X, \mathscr{A}, \mu)$. Then for any subset $B \subset A$, if B is measurable, then $\mu(B) = 0 \implies B$ is also a measure zero set.* ***But*** *we aren't guaranteed B is always measurable for every σ-algebra.*

Example 7.33. *Let $\mathscr{A} = \{\emptyset, X\}$ and $\mu = 0$. Then*

$$\mu(X) = 0$$

but for an arbitrary set $A \subset X$, where $\emptyset \neq A \neq X$, then we cannot speak to the measure of this set.

This fact may cause troubles in the integration theory, so we must overcome it by considering the completion of this measure space.

Definition 7.34. *A measure space $(X, \mathscr{A}, \mu)$ is <u>complete</u> if every subset of a measure zero set is measurable.*

Definition 7.35. *Given a measure space $(X, \mathscr{A}, \mu)$, we define $\overline{\mathscr{A}}$ to be the σ-algebra generated by*

$$\mathscr{A} \bigcup \{\text{subsets of measure zero sets}\}$$

Then

$$\overline{\mathscr{A}} = \{A : \exists E, F \in \mathscr{A} \text{ such that } E \subset A \subset F, \mu(F \setminus F) = 0\}$$

So for all $A \in \bar{\mathscr{A}}$, define

$$\overline{\mu(A)} := \mu(E) = \mu(F)$$

Then the complete measure space $(X, \overline{\mathscr{A}}, \overline{\mu})$ is called the <u>completion</u> of $(X, \mathscr{A}, \mu)$.

Example 7.36. *Given $(\mathbb{R}^n, \mathscr{B}(X)^n, \lambda)$ is not complete. It's completion is $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), \lambda)$, where $\mathcal{L}(\mathbb{R}^n)$ is the family of all Lebesgue measurable sets.*

Proposition 7.37. *Let μ be translationally invariant and locally finite measure of $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. Then $\mu = c\lambda$ for some number $c \in \mathbb{R}$. Moreover, $c = \mu([0, 1))$.*

Proof. (Idea) It is sufficient to show that

$$\mu([0, t)) = ct$$

for every $t \geq 0$. Notice

$$[0, t) = \bigcup_{i=1}^{q-1} \left[\frac{i}{q}, \frac{i+1}{q}\right)$$

for some equal division of the unit square by q-subintervals. Then

$$\implies \mu([0, 1)) = q\mu([0, \frac{1}{q}))$$
$$\implies \mu([0, \frac{1}{q})) = \frac{c}{q}$$
$$\implies \mu([0, \frac{p}{q})) = \frac{p}{q}c$$
$$\implies \mu([0, t)) = ct$$

by density of the rationals forcing a continuous measure. ■

Question 7.38. *Recall, $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), \lambda)$ is the completion of $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n), \lambda)$. Are there any measures sets in the first space that aren't measurable in the second space? Yes, Vitali's example given in the motivation is such a set.*

Theorem 7.39. *A subset $A \subset \mathbb{R}^n$ is Lebesgue measurable if and only if for every $\epsilon > 0, \exists$ a closed set F and an open set G such that*

$$F \subset A \subset G$$

and

$$\lambda(G \setminus F) \leq \epsilon$$

Moreover,

$$\begin{aligned} \lambda(A) &= \inf\{\lambda(U) : U \text{ is open }, U \supset A\} && \textit{outer regularity} \\ &= \sup\{\lambda(K) : K \text{ is closed }, K \subset A\} && \textit{inner regularity} \end{aligned}$$

Proof. Let $\epsilon = \frac{1}{k}$. Then $\exists \underbrace{G_k}_{\text{open}} \supset A \supset \underbrace{F_k}_{\text{closed}}$ where

$$\lambda(F_k \setminus G_k) \leq \frac{1}{k}$$

So

$$\underbrace{\bigcup_k G_k}_{G_\delta} \supset A \supset \underbrace{\bigcap_k F_k}_{F_\sigma}$$

∎

Theorem 7.40. *If A is Lebesgue measurable, then $\exists a, G_\delta$ set G and an F_σ set F such that $G \supset A \supset F$ and $\lambda(G \setminus F) = 0$.*

Definition 7.41. *A property that hold except on a set of measure zero is said to hold almost everywhere or a.e. for short. If there measure is obvious, then explicitly write $\mu - a.e.$.*

Example 7.42. *Two functions $f = g$ μ-a.e. means that*

$$\mu\left(\{x : f(x) \neq g(x)\}\right) = 0$$

That is, $f = g$ on $X \setminus N$ where $\mu(N) = 0$. Further,

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{Q}^c \end{cases}$$

Then $f = 0$ λ-a.e.

Example 7.43. *Suppose $f : X \to [-\infty, \infty]$. Then $|f(x)| < \infty$ μ-a.e. means*

$$\mu(\{x : |f(x)| = \infty\}) = 0$$

Example 7.44. *Let $A \subset \mathbb{R}$, define the essential supremum of A by*

$$\textit{ess sup } A = \inf\{c : x \leq c \text{ a.e. } x \in A\}$$

Further, if we let $A = (\mathbb{Q}^c \cap [0, 1]) \cup \mathbb{Z}$, then

$$\sup(A) = \infty \qquad \textit{but} \qquad \textit{ess sup } (A) = 1$$

7.3 Properties of Real-Valued Measurable Functions

Example 7.45. *A function f on $[a, b]$ is Riemann integrable if and only if f is bounded and continuous λ-a.e. on $[a, b]$.*

Lebesgue integration theory applies to more general functions called measurable functions.

Definition 7.46. *Let $(X, \mathscr{A})$ be a measurable space. A real-valued function $f : X \to \mathbb{R}$ is a measurable function (with respect to $\mathscr{A}$) if for all $t \in \mathbb{R}$,*

$$f^{-1}((t, \infty)) = \{x \in X : f(x) > t\}$$

is measurable.

Remark 7.47. *Instead of $>$, we could have chosen $\geq, \leq, <$. All these definitions are equivalent. Indeed,*

$$\{x : f(x) \geq t\} = \bigcap_n \left\{x : f(x) > t - \frac{1}{n}\right\}$$

and

$$\{x : f(x) > t\} = \bigcup_n \left\{x : f(x) \geq t + \frac{1}{n}\right\}$$

Remark 7.48. *If μ is a measure on $\mathscr{A}$, we may say f is μ-measurable. However, note that measurability does not require a measure.*

Remark 7.49. *One can also show that for any Borel set $A \subset \mathbb{R}$, the set $f^{-1}(A)$ is also σ-measurable whenever f is σ-measurable.*

Question 7.50. *In general, Let $(X, \mathscr{A}_X)$ and $(Y, \mathscr{A}_Y)$ be two measurable spaces, and $f : X \to Y$ be a function. Then $f^{-1}(\mathscr{A}_Y)$ is a σ-algebra on X. But what is the relationship between these $f^{-1}(\mathscr{A}_Y)$ and $\mathscr{A}_X$?*

Definition 7.51. *A function $f : (X, \mathscr{A}_X) \to (Y, \mathscr{A}_Y)$ between two measurable spaces is called $\underline{(\mathscr{A}_X, \mathscr{A}_Y)\text{-measurable}}$ if $f^{-1}(\mathscr{A}_Y) \subset \mathscr{A}_X$.*

Proposition 7.52. *If $\mathscr{A}_Y$ is a σ-algebra generated by a collection $\mathcal{C} \subset \mathscr{A}_Y$, then $f : (X, \mathscr{A}_X) \to (Y, \mathscr{A}_Y)$ is measurable if and only if*

$$f^{-1}(\mathcal{C}) \subset \mathscr{A}_X$$

That is, for every $E \in \mathcal{C}, f^{-1}(E) \in \mathscr{A}_X$.

Example 7.53. *Suppose $f : X \to Y$ is a continuous map between metric spaces and $\mathscr{A} = \mathscr{B}(X), \mathscr{A}_X = \mathscr{B}(Y)$. Then f is measurable because f^{-1}(open sets) is still open.*

Example 7.54. *Let $f : I \to \mathbb{R}$ be a monotonic function on the interval I. Then*

$$f^{-1}((t, \infty))$$

is either $\emptyset$ or a singleton or an interval. Therefore, $f^{-1}((t, \infty)) \in \mathscr{B}(\mathbb{R})$. That is, Borel-measurable.

Lemma 7.55. *Suppose f, g are measurable functions. Then*

$$af + bg, f \cdot g, |f|, \max\{f(x), g(x)\}, \min\{f(x), g(x)\}$$

as well as for any $\phi : \mathbb{C} \to \mathbb{C}$ continuous,

$$\phi \circ f$$

are all measurable functions. In particular,

$$f_+ = \max\{f(x), 0\}$$

$$f_- = \max\{-f(x), 0\}$$

are measurable functions.

Proof.

- To show that the max function is measurable, observe

$$\{x \in X : \max\{f(x), g(x)\} \leq t\} = \{x \in X : f(x) \leq t\} \cap \{x \in X : g(x) \leq t\}$$

which is the intersection of measurable sets.

- Observe, $|f| = \max\{f, -f\} \implies$ measurable.
- Observe, $f(x) = f_+ - f_-$ and $|f| = f_+ + f_-$ and therefore we can write the positive and negative parts of f as a linear combination of measurable functions.

■

Lemma 7.56. *If $\{f_n\}$ is a sequence of measurable functions. Then*

$$\sup_n f_n(x), \inf_n f_n(x)$$

$$\limsup_{n\to\infty} f_n(x), \liminf_{n\to\infty} f_n(x)$$

are also measurable. In particular, $f_n(x) \to f(x)$, then f is also measurable.

Proof.

$$\{x \in X : \sup_n f_n(x) \leq t\} = \bigcap_n \{x \in X : f_n(x) \leq t\}$$

$$\{x \in X : \inf_n f_n(x) < t\} = \bigcup_n \{x \in X : f_n(x) < t\}$$

which show $\sup_n, \inf_n$ are measurable. Also

$$\limsup_n f_n = \inf_k \left(\sup_{n\geq k} f_n\right)$$

is also measurable. ■

Proposition 7.57. *Suppose $(X, \mathscr{A}, \mu)$ is a complete measure space and $f = g$ μ-a.e. on X. Then f is measurable if and only if g is measurable.*

Example 7.58. *Let $(X, \mathscr{A})$ be a measurable space and $E \subset X$. Let*

$$\mathcal{X}_E(x) = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}$$

be the characteristic fucntion of E. Then $\mathcal{X}_E$ is $\mathscr{A}$-measurable if and only if $E \in \mathscr{A}$

Proof.

$$\mathcal{X}_E^{-1}((t, \infty)) = \emptyset, E, \text{ or } X$$

■

Example 7.59. *Let $A_1, A_2, \ldots, A_n$ be measurable sets in X and $c_1, \ldots, c_n \in \mathbb{R}$. Then*

$$\phi(x) = \sum_{i=1}^n c_i \mathcal{X}_{A_i}(x)$$

is still measurable. Such functions are called simple functions.

Remark 7.60. *The representation of a simple function as a sum of characteristic functions is not unique. The standard representation uses disjoint sets $\{A_i\}$ and distinct, non-zero c_i's. How?: A simple function ϕ can only take finitely many distinct and non-zero values: $c_1, \ldots, c_n$. Then, set each A_i to be its preimage*

$$A_i = \phi^{-1}(c_i)$$

Theorem 7.61 (Approximation by Simple Functions)**.** *Let f be a non-negative measurable function on X. Then, there exists a monotone increasing sequence of non-negative simple functions*

$$0 \leq \phi_1(x) \leq \phi_2(x) \leq \ldots$$

such that

$$\lim_{k\to\infty} \phi_k(x) = f(x)$$

for every $x \in X$. Moreover, if f is bounded, then the convergence is is uniform.

Proof. (Idea) Subdivide the range of f. Then for all $n = 0, 1, 2, \ldots$, define

$$\phi_n(x) = \begin{cases} \frac{i-1}{2^n} & \frac{i-1}{2^n} \leq f(x) < \frac{i}{2^n},\ \text{for} i = 1, 2, 3, \ldots, 2^n \\ 2^n & \text{if } f(x) \geq 2^n \end{cases}$$

Then $\{\phi_n(x)\}$ is an increasing sequence and

$$0 \leq f - \phi_n \leq \frac{1}{2^n} \text{ if } f(x) < 2^n$$

$$0 \leq f - \phi_n \leq f - 2^n \text{ if } f(x) > 2^n$$

■

Proposition 7.62. *Let f be a measurable function on X. Then, there exists a sequence of simple functions $\{\phi_k(x)\}$ such that*

$$|\phi_k(x)| \leq f(x) \text{ and } \lim_k \phi_k(x) = f(x)$$

for all $x \in X$. Then convergence is uniform if f is bounded.

8 Integration Theory

Recall the function $f = \mathcal{X}_{\mathbb{Q}\cap[0,1]}$ is not Riemann integrable but $f = 0$ λ-a.e. We hope to construct an integration theory such that

$$\int_0^1 f d\lambda = \int_0^1 0 d\lambda = 0$$

in some sense. This will lead us to Lebesgue integration.
We want $f = g$ μ-a.e. then $\int f d\mu = \int g d\mu$. How to define it?

Definition 8.1. *1. For Simple Functions: Consider ϕ of the form*

$$\phi = \sum_{i=1}^{n} c_i \mathcal{X}_{A_i}$$

Then we define

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

Remark 8.2. *This value is independent of how ϕ is represented. Usually, we use the standard representation.*

2. *Non-negative Measurable functions: Let $f : X \to [0, \infty]$ be a non-negative measurable function on a measurable space $(X, \mathscr{A}, \mu)$. Define*

$$\int_X f d\mu = \sup \left\{ \int_X \phi d\mu : \phi \textit{ is simple and } 0 \leq \phi \leq f \right\}$$

If $\int_X f d\mu < \infty$, then we say that f is μ-integrable.

Remark 8.3. *There is another, but equivalent using Riemann integration:*

$$\int_X f d\mu = \int_0^\infty \mu(f^{-1}(t, \infty))\ dt$$

Why? Because as $t \to \infty$, $\mu(f^{-1}(t, \infty))$ is a monotonically non-increasing function of t and therefore is Riemann integrable. Moreover, the two must be consistent.

Example 8.4.

$$\int c\mathcal{X}_A\ d\mu = \int_0^c \mu(A)\ dt = c\mu(A)$$

Question 8.5. *How is Riemann-integration different from Lebesgue integration? Riemann integration partitions the area over the x-axis, while Lebesgue integration partitions the area over the y-axis!*

3. *General Case: If $f : X \to \overline{\mathbb{R}} = [-\infty, \infty]$ is measurable. The $f = f_+ - f_-$. So we can define:*

$$\int_X f d\mu = \int_X f_+\ d\mu - \int_X f_- d\mu$$

If $\int_X |f|\ d\mu < \infty$, if both $\int_X f_\pm d\mu < \infty$, then we say that f is μ-integrable.

Remark 8.6. *f integrable $\iff$ both f_+ and f_- are integrable $\iff$ $|f|$ is integrable*

Example 8.7. *For every $A \in \mathscr{A}$, define*

$$\int_A f d\mu = \int_X f\mathcal{X}_A\ d\mu$$

Example 8.8. *$\mu = \delta_{x_0}$ with $x_0 \in X$, then*

$$\int \mathcal{X}_A d\mu = \mu(A) = \delta_{x_0}(A) = \mathcal{X}_A(x_0)$$

Further, if $\phi \sum_{i=1}^n c_i \mathcal{X}_{A_i}$, then

$$\int \phi \; d\mu = \sum_{i=1}^n c_i \mu(A_i) \sum_{i=1}^n c_i \mathcal{X}_{A_i}(x_0) = \phi(x_0)$$

then for any measurable function,

$$\int f d\mu = f(x_0)$$

Example 8.9. *Let μ be the counting measure on $X = \mathbb{N}$ where $f : \mathbb{N} \to \mathbb{R}$, then*

$$\int f \; d\mu = \sum_{n=1}^{\infty} f_n$$

and f is μ-integrable $\iff$ the series $\sum_n f_n$ is absolutely convergent.

Lemma 8.10 (Properties of Integral Functions)**.** *Suppose f, g are μ-integrable. Then*

1. *Linearity:*
$$\int (af + bg) \; d\mu = a \int f d\mu + b \int g d\mu$$
2. *Additivity: If E and F are disjoint, measurable sets, then*
$$\int_{E \cup F} f \; d\mu = \int_E f \; d\mu + \int_F f \; d\mu$$
3. *Monotonicity: If $f \leq g$, then*
$$\int f \; d\mu \leq \int g d\mu$$
4. $$\left| \int f \; d\mu \right| \leq \int |f| d\mu$$
5. *$f : X \to [-\infty, \infty]$ is μ-integrable $\implies$ $f(x)$ is finite μ-a.e. (Might be asked to prove - potential test question)*

Corollary 8.10.1. *If $\int |f_n - f| \; d\mu \to 0$, then $\int f_n d\mu \to \int f d\mu$*

8.1 Littlewood's Three Principles

These provide very useful, intuitive guides about measurable sets / functions.

Theorem 8.11. 1. *Every sets is nearly a finite union of intervals. More precisely, if $E \subset \mathbb{R}^n$ is a Lebesgue measurable set with $\lambda(E) < \infty$, then for every $\epsilon > 0$, there exists a finite union*

$$F = \bigcup_{i=1}^N Q_i$$

of closed cubes such that $\lambda(E \Delta F) \leq \epsilon$

2. *(Lusin) Every function is nearly continuous. In particular, suppose f is measurable and finite-valued on E with $\lambda(E) < \infty$. Then $\forall \epsilon > 0$, there exists a closed set $F_\epsilon \subset E$ and $\lambda(E \setminus F_\epsilon) \leq \epsilon$ such that*
$$f|_{F_\epsilon}$$
is continuous. (i.e. f can only be ill-behaved on a set of measure zero).

3. *(Egorov) Every convergent sequence of functions is nearly uniformly convergent. In particular, suppose $\{f_k\}_k$ is a sequence of measurable functions defined on a measurable set E with $\mu(E) < \infty$ and assume $f_n \to f$ pointwise μ-a.e. on E. Then $\forall \epsilon > 0, \exists$ a closed set $A_\epsilon \subset E$ such that*
$$\mu(E \setminus A_\epsilon) \leq \epsilon$$
and $f_n \to f$ uniformly on A_ϵ.

8.2 Convergence Theorems

Suppose $f_n \to f$ pointwise.

Question 8.12. *When can we assert*
$$\int f_n \, d\mu \to \int f \, d\mu$$
In Riemann integration, we need $f_n \to f$ uniformly. Instead, what does Lebesgue integration require?

Theorem 8.13 (Bounded Convergence Theorem)**.** *Suppose that*

1. $\mu(X) < +\infty$
2. $\{f_n : X \to \mathbb{R}\}$ *is a measurable sequence of functions and* $|f_n(x)| \leq M$ *for all* $n \in \mathbb{N}, x \in \mathbb{R}$.
3. $f_n(x) \to f$ *pointwise* μ*-a.e.*

Then f is measurable, bounded and
$$\int_X |f_n - f| \, d\mu < +\infty$$
Moreover, $\int_X f_n \to \int_X f$

Proof. For all $\epsilon > 0$, by Egorov's theorem, there must exist a set $A_\epsilon \subset X$ such that $f_n \rightrightarrows f$ uniformly on A_ϵ and $\mu(X \setminus A_\epsilon) \leq \epsilon$. So
$$\begin{aligned}\int_X |f_n - f| d\mu = \int_{A_\epsilon} |f_n - f| d\mu + \int_{X \setminus A_\epsilon} |f_n - f| \, d\mu \\ &\leq \epsilon\mu(A_\epsilon) + 2M\mu(X \setminus A_\epsilon) \\ &\leq \epsilon\mu(X) + 2M\epsilon \to 0\end{aligned}$$
∎

Question 8.14. *What happens if $\mu(X) = \infty$ or $\{f_n\}$ is not uniformly bounded. Two cases:*

1. *Mass leaks to infinity: Define*
$$f_n = \begin{cases} 1 & x \in [n, n+1] \\ 0 & \text{otherwise} \end{cases}$$
$f_n \to f = 0 \ \forall x$ *but* $\int_{\mathbb{R}} f_n \, dx = 1 > 0 = \int_{\mathbb{R}} f \, dx$

2. *A point possesses infinite mass: Let f_n be defined such that*

$$f_n(x) = \begin{cases} n & x = 0 \\ 0 & |x| > \frac{1}{2n} \\ \text{linearly continuous} & \text{everywhere else} \end{cases}$$

Then $f_n \to f = 0$ but

$$\int_{\mathbb{R}} f_n dx = 1 \neq 0 = \int_{\mathbb{R}} f \ dx$$

Nevertheless, we have the following important theorem that provides a theoretical bounding of these situations:

Lemma 8.15 (Fatou)**.** *Suppose $\{f_n\}_n$ is a sequence of non-negative functions. Then*

$$\int_X \liminf_{n\to\infty} f_n(x) \ d\mu \leq \liminf_{n\to\infty} \int_X f_n \ d\mu$$

Proof. Let $f(x) = \liminf_{n\to\infty} f_n(x)$. It suffices to show $\int_X g \ d\mu \leq \liminf_{n\to\infty} f_n d\mu$. Observe, for all $0 \leq g \leq f$ such that g is bounded and has support on a set E of finite measure. Doing this, we can use bounded convergence theorem since f is not bounded, but f is. Take

$$g_n = \min\{g, f_n\}$$

Then $g_n \to g$ a.e. and g_n is supported on E. By bounded convergence theorem, we have

$$\int_X g_n \ d\mu \to \int_X g d\mu$$

but $g_n \leq f_n \implies \int g_n d\mu \leq \int f_n d\mu$. So

$$\int f d\mu \leq \liminf_{n\to\infty} f_n \ d\mu$$

∎

Corollary 8.15.1. *Suppose $\{f_n\}$ and f are non-negative measurable functions with $f_n(x) \to f(x)$ μ-a.e. If $f_n(x) \leq f(x)$ for all n, then*

$$\liminf_{n\to\infty} \int_X f_n \ d\mu = \int_X f \ d\mu$$

Proof.

$$f_n \leq f \implies \int f_n d\mu \leq \int f d\mu$$

$$\implies \overline{\lim_{n\to\infty}} \int f_n \leq \int f d\mu = \int \underline{\lim_{n\to\infty}} f_n \ d\mu \leq \underline{\lim_{n\to\infty}} \int f_n$$

∎

Corollary 8.15.2 (Monotone Convergence Theorem)**.** *Suppose $\{f_n\}$ is a monotonically increasing sequence of non-negative measurable functions with $f_n \to f$ a.e. and*

$$\lim_n f_n(x) = f(x)$$

Then

$$\int_X f_n d\mu \to \int_X f d\mu$$

Example 8.16. *Let $f_n(x) = \sum_{k=1}^{n} a_k(x)$ with $a_k(x) \geq 0$ is measurable for each k. Then*

$$f_n(x) = \sum_{k=1}^{n} a_k(x) \to \sum_{k=1}^{\infty} a_k(x) = f(x).$$

By monotone convergence theorem, we know

$$\lim_n \int_X f_n \, d\mu = \int_X f(x) d\mu$$

$$\implies \lim_n \int_X \sum_{k=1}^{n} a_k(x) d\mu = \sum_{k=1}^{\infty} \int_X a_k(x) d\mu$$

That is,

$$\sum_{k=1}^{\infty} \int_X a_k(x) d\mu = \int_X \sum_{k=1}^{\infty} a_k(x) d\mu$$

This is called Beppo Levi's Theorem.

Note 8.17. *If $\sum_{k=1}^{\infty} \int_X a_k(x) d\ mu < \infty$, then $f(x) = \sum_{k=1}^{\infty} a_k(x)$ is μ-integrable. Thus, $f(x)$ is finite μ-a.e. That is, the series*

$$\sum_{k=1}^{\infty} a_k(x)$$

converges for μ-a.e.

Now we arrive at the cornerstone of Lebesgue integration theory:

Theorem 8.18 (Lebesgue Dominated Convergence)**.** *Suppose the $\{f_n\}$ is a sequence of measurable functions on a measure space $(X, \mathscr{A}, \mu)$ such that $f_n(x) \to f(x)$ pointwise μ-a.e.*
If there is a non-negative μ-integrable function $g : X \to [0, \infty]$ such that

$$|f_n(x)| \leq g(x) \quad \forall n \in \mathbb{N}, x \in X$$

Then

$$\lim_n \int_X f_n d\mu = \int_X f d\mu$$

Proof. Letting $n \to \infty$, we see $|f(x)| \leq g(x)$. Also, since g is μ-integrable, f must also be μ-integrable. Since $g + f \geq 0$, then by Fatou's lemma:

$$\int_X (g + f) d\mu \leq \liminf_n \int_X (g + f_n) d\mu$$

Consequently,

$$\int_X f d\mu \leq \liminf_n \int_X f_n \, d\mu$$

Similarly,

$$\int_X (g - f) d\mu \leq \liminf_n \int_X (g - f_n) d\mu$$

$$\implies \int_X f d\mu \geq \limsup_n \int_X f_n d\mu$$

Thefore,

$$\int_X f d\mu = \lim_n \int_X f_n d\mu$$

■

8.3 Product Measures

Let $(X, \mathscr{A})$ and $(X, \mathscr{B})$ be two measurable spaces. Then

$$X \times Y = \{(x, y) : x \in X, y \in Y\}$$

$$\mathscr{A} \times \mathscr{B} := \text{ the } \sigma\text{-algebra generated by measurable rectangles}$$

Suppose that $E \subset X \times Y$. Then for all $x \in X, y \in Y$, we define:

$$x - \text{section } E_x := \{y \in Y : (x, y) \in E\}$$

$$y - \text{section } E_y := \{x \in X : (x, y) \in E\}$$

Proposition 8.19. *For any $E \subset \mathscr{A} \times \mathscr{B}$, it holds that $E_x \in \mathscr{B}$ and $E_y \in \mathscr{A}$ for every $x \in X, y \in Y$.*

Proof. Let $M := \{E \subset X \times Y : E_x \in \mathscr{B}, E_y \in \mathscr{A} \ \forall \ (x, y) \in X \times Y\}$ Then M contains all measurable rectangles and M is a σ-algebra. Therefore $\mathscr{A} \times \mathscr{B} \subset M$. ■

Question 8.20. *Now, given two measure spaces $(X, \mathscr{A}, \mu)$ and $(Y, \mathscr{B}, \nu)$, how can we define a measure $\mu \times \nu$ on $\mathscr{A} \times \mathscr{B}$?*

Approach #1: $\int_X \nu(E_x) d\mu(x)$
Approach #2: $\int_Y \mu(E_y) d\nu(y)$

Question 8.21. *Are these two compatible / equal? Generally, no, these are not equal.*

Example 8.22. *Let $X = Y = [0, 1]$ with $\mu =$ Lebesgue measure, $\nu =$ counting measure. Define*

$$E = \{(x, y) : x = y\}$$

Then

$$\int_X \nu(E_x) d\mu(x) = \int_X 1 d\mu(x) = 1$$

$$\int_Y \mu(E_y) d\nu(y) = \int_Y 0 d\nu(y) = 0$$

Fix: The two approaches are compatible when both μ and ν are σ-finite on the measure space!

Definition 8.23. *Let $(X, \mathscr{A}, \mu)$ and $(Y, \mathscr{B}, \nu)$ be two σ-finite measure spaces. For any $E \in \mathscr{A} \times \mathscr{B}$, define*

$$(\mu \times \nu)(E) := \int_X \nu(E_x) d\mu(x) = \int_Y \mu(E_y) d\nu(y)$$

$\mu \times \nu$ is a measure on $\mathscr{A} \times \mathscr{B}$, called the product measure.

Theorem 8.24 (Fubini-Tonelli Theorem)**.** *Suppose $(X, \mathscr{A}, \mu)$ and $(Y, \mathscr{B}, \nu)$ are two σ-finite measure spaces, and*

$$f : X \times Y \to [-\infty, \infty]$$

is a $\mathscr{A} \times \mathscr{B}$ function. Then

1. *(Tonelli) If $0 \leq f \leq \infty$, then*

$$\int_{X \times Y} f(x, y) d(\mu \times \nu) = \int_X \left(\int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X f(x, y) d\mu(x) \right) d\nu(y)$$

That is, iterated integrals are the same as double integrals.

Note 8.25. *$\int_Y f(x,y)\ d\nu(y)$ is a $\mathscr{A}$-measurable function of x, and $\int_X f(x,y)\ d\mu(x)$ is a $\mathscr{B}$-measurable function of y.*

2. *(Fubini) If f is integrable, then*

$$\int_{X\times Y} f(x,y)d(\mu\times\nu) = \int_X\left(\int_Y f(x,y)d\nu(y)\right)d\mu(x) = \int_Y\left(\int_X f(x,y)d\mu(x)\right)d\nu(y)$$

Note 8.26. *Here, for μ a.e. $x\in X, f(x,y)$ is integrable on Y and $\int Y f(x,y)d\nu(y)$ is μ-integrable on X. For ν a.e. $y\in Y, f(x,y)$ is integrable on X and $\int X f(x,y)d\mu(x)$ is ν-integrable on Y.*

Remark 8.27. *In practice, given a measurable function f on $X\times Y$, if we want to calculate $\int_{X\times Y} f(x,y)\ d(\mu\times\nu)$, we consider the following steps:*

1. *Calculate one of the iterated integrals of $|f|$ freely,*

$$\int_X\left(\int_Y |f|\right) \text{ or } \int_Y\left(\int_X |f|\right)$$

2. *If the above integral is finite, the by Tonelli's theorem, we know*

$$\int_{X\times Y} f(x,y)\ d(\mu\times\nu) < \infty \implies f \text{ integrable}$$

3. *Apply Fubini's theorem,*

$$\implies \int_{X\times Y} f(x,y)d(\mu\times\nu) = \int_X\left(\int_Y f(x,y)d\nu(y)\right)d\mu(x) = \int_Y\left(\int_X f(x,y)d\mu(x)\right)d\nu(y)$$

Remark 8.28. *If $\int_{X\times Y} f(x,y)\ d(\mu\times\nu)=\infty$, then the two iterated integrals by have different finite values.*

Example 8.29.

$$f(x,y) = \frac{x^2-y^2}{(x^2+y^2)^2} \text{ on } [0,1]^2$$

$$\implies \int_0^1 f(x,y)\ dy = \frac{1}{1+x^2} \implies \int_0^1\left(\int_0^1 f(x,y)\ dy\right)dx = \frac{\pi}{4}$$
$$\implies \int_0^1 f(x,y)\ dx = \frac{-1}{1+y^2} \implies \int_0^1\left(\int_0^1 f(x,y)\ dx\right)dy = \frac{-\pi}{4}$$

Example 8.30. *Let $\mu=\nu$ be the counting measure on $\mathbb{N} \implies X=Y=\mathbb{N}$. Then, by Fubini-Tonelli's Theorem,*

1. *If $a_{mn}\geq 0$, then*

$$\sum_{m,n=1}^{\infty} a_{m,n} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{m,n} = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{m,n}$$

2. *Suppose $a_{m,n}\in\mathbb{R}$ or $\mathbb{C}$. If $\sum_{m,n=1}^{\infty}|a_{m,n}|<+\infty$, then*

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{m,n} = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{m,n}$$

Example 8.31. *Calculate Gaussian integral: Let* $I = \int_{-\infty}^{\infty} e^{-x^2}\ dx$, *then*

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2}\ dx\right)\left(\int_{-\infty}^{\infty} e^{-y^2}\ dy\right) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2-y^2}\ dx\ dy \\ &= \int_0^{2\pi}\int_0^{\infty} e^{-r^2} r\ dr\ d\theta \\ &= \int_0^{2\pi} \frac{1}{2}\ d\theta = \pi \end{aligned}$$

$$\implies \int_{-\infty}^{\infty} e^{-x^2}\ dx = \sqrt{\pi}$$

Example 8.32 (Layer Cake Representation)**.** *Let f be measurable, with $1 \leq p < \infty$. We have*

$$\int_X |f|^p\ d\mu = p\int_0^{\infty} t^{p-1}\mu(\{x : |f(x)| > t\})\ dt$$

$$p = 1 \implies \int_X |f|\ d\mu = \int_0^{\infty} \mu(\{x : |f(x)| > t\})\ dt$$

Why is this true? Observe:

$$\begin{aligned} \int_X |f|^p\ dx &= \int_X \left(\int_0^{|f(x)|} pt^{p-1}\ dt\right) d\mu(x) \\ &= \int_X \left(\int_0^{\infty} pt^{p-1}\mathcal{X}_{[0,|f(x)|]}(t)\ dt\right) d\mu(x) \\ &= \int_0^{\infty} \left(\int_X pt^{p-1}\mathcal{X}_{[0,|f(x)|]}(t)\ d\mu(x)\right)\ dt && \textit{Tonelli} \\ &= \int_0^{\infty} pt^{p-1}\left(\int_X \mathcal{X}_{\{x:|f(x)|>t\}}\right) dt \\ &= p\int_0^{\infty} t^{p-1}\mu(\{x : |f(x)| > t\})\ dt \end{aligned}$$

Example 8.33.

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

Suppose $f \geq 0, g \geq 0$ are both integrable. Then

$$\begin{aligned} \int_{\mathbb{R}^n} f * g(x)\ dx &= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} f(x-y)g(y)\ dy\right)\ dx \\ &= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} f(x-y)g(y)\ dx\right)\ dy && \textit{Tonelli} \\ &= \int_{\mathbb{R}^n} g(y)\left(\int_{\mathbb{R}^n} f(x-y)\ dx\right)\ dy \\ &= \left(\int_{\mathbb{R}^n} f(x)\ dx\right)\left(\int_{\mathbb{R}^n} g(y)\ dy\right) && \textit{Change of variables} \end{aligned}$$

$$\implies \|f * g\|_1 = \|f\|_1\,\|g\|_1$$

9 The L^p Spaces

Let $(X, \mathscr{A}, \mu)$ be a measure space with $1 \leq p < \infty$. Two functions $f, g : X \to \mathbb{R}$, (or $\mathbb{C}$) is equivalent if $f = g$ μ-a.e.

Definition 9.1. *Let*

$$L^p(X, \mu) := \{f : X \to \mathbb{R} : f \text{ is measurable and } \int_X |f|^p \, d\mu < \infty\}$$

with respect to the equivalence relation of a.e. equality.

Definition 9.2. *We equip $L^p(x, \mu)$ with the* *L^p norm of f*, *defined by*

$$\|f\|_p = \left(\int_X |f|^p \, d\mu\right)^{\frac{1}{p}}$$

Definition 9.3. *$L^\infty(X, \mu)$ consists of equivalence classes of bounded functions $f : X \to \mathbb{R}$ (or $\mathbb{C}$) such that $\exists M \geq 0$ such that*

$$|f(x)| \leq M \quad \mu - a.e.$$

The L^∞ norm of f is defined by

$$\|f\|_\infty := \inf\{M : |f(x)| \leq M, \mu - a.e.\}$$

This norm is often referred to as the essential supremum of f.

9.1 Some Inequalities

Lemma 9.4 (Hölder)**.** *For all $f \in L^p(X, \mu), g \in L^q(X, \mu)$ where $\frac{1}{p} + \frac{1}{q} = 1$ and $1 \leq p \leq \infty$, then*

$$\|f \cdot g\|_1 \leq \|f\|_p \|g\|_q$$

Lemma 9.5 (Generalized Hölder)**.** *If $1 \leq p_i \leq \infty$ for $i = 1, 2, \ldots, n$ which satisfies*

$$\sum_{i=1}^n \frac{1}{p_i} = 1$$

and $f_i \in L^{p_i}(X, \mu)$ for all i, then

$$f_1 \cdot f_2 \cdot \ldots \cdot f_n \in L^1(X, \mu)$$

and

$$\|f_1 \cdot f_2 \cdot \ldots \cdot f_n\|_1 \leq \|f_1\|_{p_1} \cdot \|f_2\|_{p_2} \cdot \ldots \cdot \|f_n\|_{p_n}$$

Lemma 9.6 (Minkowski)**.** *For all $f, g \in L^p, 1 \leq p \leq \infty$,*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

Lemma 9.7 (Young)**.** *Suppose that $1 \leq p, q \leq \infty$ satisfies*

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

If $f \in L^p(\mathbb{R}^n), g \in L^q(\mathbb{R}^n)$, then $f * g \in L^r(\mathbb{R}^n)$ with

$$\|f * g\|_r \leq \|f\|_p \|g\|_q$$

Theorem 9.8. *If $(X, \mathscr{A}, \mu)$ is a measure space, then $L^p(X, \mu)$ is Banach for $1 \leq p \leq \infty$.*

Proof. **Idea:** Only need to prove L^p is Cauchy-complete. So we first consider $\{f_n\}$ be a Cauchy sequence in L^p. We then prove that it converges by the following steps:

1. We want to construct a candidate f such that $f_{i_k} \to f$ pointwise μ-a.e.
   For $\epsilon = \frac{1}{2}, \exists i_1$ such that
   $$\|f_{i_1} - f_n\| \leq \frac{1}{2} \text{ for all } n \geq i_1$$
   For $\epsilon = \frac{1}{4}, \exists i_2$ such that
   $$\|f_{i_1} - f_n\| \leq \frac{1}{4} \text{ for all } n \geq i_2$$
   and so on. So we get a subsequence $\{f_{i_k}\}$ with the property that
   $$\left\|f_{i_k} - f_{i_{k+1}}\right\| \leq \frac{1}{2^k}, \quad k \in \mathbb{N}$$
   Claim: $\{f_{i_k}\}$ converges pointwise μ-a.e.
   Let $F_n(X) = |f_i(x)| + \sum_{k=1}^n |f_{i_k}(x) - f_{i_{k+1}}(x)|$. Then $F_n(x)$ is a monotone sequence of non-negative functions. Then
   $$F(x) = \lim_{n\to\infty} F_n(x) \in [0, \infty] \text{ for every } x \in X$$
   Notice: by Minkowski
   $$\begin{aligned}\|F_n\|_p &\leq \|f_{i_1}\|_p + \sum_{k=1}^n \left\|f_{i_k} - f_{i_{k+1}}\right\|_p \\ &\leq \|f_{i_1}\|_p + 1\end{aligned}$$
   By Monotone Convergence Theorem , because $\|F\|_p \leq \|f_{i_1}\|_p + 1$, we conclude $F \in L^p(X)$ and F is finite μ-a.e. Therefore, the sequence
   $$f_{i_{n+1}} = f_{i_1} + \sum_{k=1}^n (f_{i_{k+1}}(x) - f_{i_k}(x))$$
   converges absolutely for μ-a.e. x and hence it converges for the same x's to some number $f(x)$.

2. *Claim:* $f_{i_k} \to f$ in L^p.
   Since
   $$|f_{i_{n+1}}(x)| \leq |f_{i_1}(x)| + \sum_{k=1}^n |f_{i_{k+1}}(x) - f_{i_k}(x)| = F_n(x) \leq F(x)$$
   and $F \in L^p$, by the Lebesgue Dominated Convergence Theorem, $f \in L^p(X, \mu)$. Again, since
   $$|f_{i_n}(x) - f(x)| \leq F(x) + |f(x)| \in L^p(X, \mu)$$
   By Lebesgue Dominated Convergence Theorem again,
   $$\lim_{n\to\infty} \|f_{i_n} - f\|_p = \left(\int_X \lim_{n\to\infty} |f_{i_n} - f|^p \, d\mu\right)^{\frac{1}{p}} = (0 \, d\mu)^{\frac{1}{p}} = 0$$
   Therefore,
   $$f_{i_n} \to f \text{ in } L^p$$

3. Since $\{f_n\}$ is Cauchy, $f_{i_n} \to f$ in L^p we have $f_n \to f$ in L^p.

■

Theorem 9.9. *If $\mu(x) < \infty, 1 \le p < q < \infty$, then $L^p \supset L^q$. Moreover,*

$$L^1(X,\mu) \subset L^2(X,\mu) \ldots \supset L^\infty(X,\mu)$$

Proof. Define $r \in [1,\infty]$ by

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

So for all $f \in L^q(X,\mu)$

$$\begin{aligned}
\|f\|_p^p &= \int_X |f|^p \, d\mu \\
&= \int_X |f|^p \cdot 1 \, d\mu \\
&\le \|f^p\|_{\frac{q}{p}} \|1\|_{\frac{r}{p}} && \text{Hölder Inequality} \\
&= \left(\int |f|^q d\mu\right)^{\frac{p}{q}} \cdot \left(\int_X d\mu\right)^{\frac{p}{r}} \\
&= \|f\|_q^p \, \mu(X)^{\frac{p}{r}}
\end{aligned}$$

Therefore,

$$\|f\|_p \le \|f\|_q \cdot \mu(X)^{\frac{1}{r}} < \infty$$

■

Example 9.10. *Let $X = (0,1)$ and $f(x) = x^r$ with $\mu =$ Lebesgue Measure. Then*

$$f \in L^p(X) \iff \int_0^1 |f|^p \, d\mu < \infty \iff \int_0^1 x^{pr} dx < \infty \iff pr > -1$$

Example 9.11.

$$\frac{1}{\sqrt{x}} = x^{\frac{-1}{2}} \in L^1(X) \text{ but } \frac{1}{\sqrt{x}} \notin L^2(X) \implies L^2(X) \text{ proper} \subset L^1(X)$$

Example 9.12.

$$\ln(\frac{1}{x}) \in L^p(X) \text{ but } \ln(\frac{1}{x}) \notin L^\infty(X)$$

Example 9.13. *Let $X = (1,\infty), f(x) = x^r$. Then*

$$f \in L^p(X) \iff pr < -1$$

$$\implies \frac{1}{x} \in L^2(X) \text{ but } \frac{1}{x} \notin L^1(X)$$

So

$$L^2(1,\infty) \not\subset L^1(1,\infty)$$

Remark 9.14. *In general, when $\mu(X) = \infty$, there is no relation of the type $L^p \subset L^q$.*

Lemma 9.15. *If $1 \le p < q < r < +\infty$, then*

- $L^q \subset L^p + L^r$
- $L^p \cap L^r \subset L^q$

9.2 Dense Subsets of $L^p(X,\mu)$

Theorem 9.16. *If $X \subset \mathbb{R}^n, 1 \le p < \infty$, then*

$$C_c(X) := \{ \text{ all continuous functions on } X \text{ with compact support}\}$$

is dense in $L^p(X,\lambda)$. That is, for all $f \in L^p(X,\lambda)$, and for all $\epsilon > 0, \exists g \in C_c(X)$ such that

$$\|f-g\|_p < \epsilon$$

Proof.

1. When $f = \mathcal{X}_A$ for some bounded measurable set $A \subset X \subset \mathbb{R}^n$ and given $\epsilon > 0$, by the Borel regularity of Lebesgue measure, there exists a bounded open set G and a compact set K such that

$$K \subset A \subset G \text{ and } \mu(G \setminus K) < \epsilon$$

Let $g \in C_c(\mathbb{R}^n)$ be a Urysohn function such that $g = 1$ on K while $g = 0$ on G^c and $0 \le g \le 1$. We know such a function exists by the example:

$$g(x) = \frac{d(x,G^c)}{d(x,K)+d(x,G^c)}$$

where the distance function is defined

$$d(x,F) = \inf\{d(x,y) : y \in F\}$$

Then

$$\begin{aligned}\|\mathcal{X}_A - g|_X\|_p &= \left(\int_{(G\setminus K)\cap X} |\mathcal{X}_A - g|^p d\lambda\right)^{\frac{1}{p}} \\ &\le \lambda(G\setminus K)^{\frac{1}{p}} < \epsilon^{\frac{1}{p}}\end{aligned}$$

2. When $f = \sum_{i=1}^N a_i\mathcal{X}_{A_i}, a_i > 0, A_i \subset X$ bounded, for each $i, \exists g_i \in C_c(X)$ such that

$$\|g_i - \mathcal{X}_{A_i}\|_p \le \frac{\epsilon}{Na_i}$$

Then

$$\begin{aligned}\left\|f - \sum_{i=1}^N a_i g_i\right\|_p &= \left\|\sum_{i=1}^N a_i(g_i - X_{A_i})\right\|_p \\ &\le \sum_{i=1}^N a_i \|g_i - X_{A_i}\|_p \\ &\le \sum_{i=1}^N a_i \cdot \frac{\epsilon}{Na_i} = \epsilon\end{aligned}$$

3. If $f \in L^p$, by homework, $\exists$ a simple function s such that

$$\|f-s\|_p < \epsilon$$

■

Remark 9.17. *$(C_c(\mathbb{R}^n, \|\cdot\|_p)$ is a metric space for $1 \le p \le \infty$. But when $1 \le p < \infty, (C_c(\mathbb{R}^n, \|\cdot\|_p)$ is not complete. Its completion is $L^p(\mathbb{R}^n)$. But when $p = \infty$, then the completion of $(C_c(\mathbb{R}^n, \|\cdot\|_p)$ is not $L^\infty(\mathbb{R}^n)$. Instead,*

$$\overline{C_c(\mathbb{R}^n} = C_0(\mathbb{R}^n) := \{ \text{ continuous functions on } \mathbb{R}^n \text{ which vanish at} \infty\}$$

Theorem 9.18. *If $1 \leq p < \infty$, then $C_c(\mathbb{R}^n$ = the space of all C^∞ functions on $\mathbb{R}^n$ with compact support is a dense subset of $L^p(\mathbb{R}^n)$.*

Proof. Let

$$\eta = \begin{cases} c \cdot \exp\left\{\frac{1}{|x|^2-1}\right\} & |x| < 1 \\ 0 & |x| \geq 1 \end{cases}$$

where the constant c is adjusted so that

$$\int_{\mathbb{R}^n} \eta(x)dx = 1.$$

Then $\eta \in C_c^\infty(\mathbb{R}^n)$. Define

$$\eta_\epsilon := \frac{1}{\epsilon^n}\eta\left(\frac{x}{\epsilon}\right)$$

We refer to η_ϵ as the standard mollifier. It satisfies

$$\int_{\mathbb{R}^n} \eta_\epsilon(x)\ dx = 1$$

and

$$supp(\eta_\epsilon) = \overline{B_\epsilon(0)}$$

Then for every $f \in L^p(\Omega), \Omega \subset \mathbb{R}^n$ open, we define

$$f^\epsilon := \eta_\epsilon * f$$

Then $f^\epsilon \in C^\infty(\Omega_\epsilon)$ where

$$\Omega_\epsilon := \{x \in \Omega : dist(x, \partial\Omega) > \epsilon\}$$

- $f \in C(\Omega) \implies f^\epsilon \rightrightarrows f$ uniformly on a compact subset.
- $f \in L^p(\Omega), 1 \leq p < \infty$, then $f^\epsilon \to f$ in $L^p(\Omega)$.

∎

10 Fourier Series

10.1 Space of Interest

Aim: Study a special Hilbert space $L^2(X)$ where X is the unit circle.

Note 10.1. *There are several ways to represent functions on the unit circle.*

- Functions on the circle $F : \mathbb{S}^1 \to \mathbb{C}$
- 2π-periodic function on $\mathbb{R}$
$$f : \mathbb{R} \to \mathbb{R} \text{ with } f(x + 2\pi) = f(x) \forall x \in \mathbb{R}$$
- Functions on the one-dimensional Torus where $\bar{T} = \mathbb{R}/2\pi\mathbb{Z}$. Here $\bar{T}$ is achieved by identifying points in $\mathbb{R}$ that differ by $2\pi n$. That is,
$$x \sim y \iff x - y = 2\pi n, n \in \mathbb{Z}$$
$$\implies [x] = \{y : x \sim y\} = \{x + 2\pi n : n \in \mathbb{Z}\} = x + 2\pi\mathbb{Z}$$
So $\bar{T} = \{[x] : x \in \mathbb{R}\}$.
- Functions on a closed interval $[a, a+2\pi]$ with $f(a) = f(a+2\pi)$. Typically, we choose $[0, 2\pi]$ or $[-\pi, \pi]$. Thus,
$$L^2(\bar{T}) = \{f : \bar{T} \to \mathbb{C} : \int_{\bar{T}} |f(x)|^2 \, dx < \infty\}$$
It is a Hilbert space with inner product
$$\langle f, g \rangle = \int_{\bar{T}} \overline{f(x)} g(x) \, dx$$

10.2 Orthonormal Basis

Let
$$e_n := \frac{1}{\sqrt{2\pi}} e^{inx}$$

Claim: $\{e_n\}_{n=1}^{\infty}$ forms an orthonormal basis for $L^2(\bar{T})$. This requires two ideas:

- We need to show this is orthonormal:
$$\begin{aligned}
\langle e_n, e_m \rangle &= \int_{\bar{T}} \frac{1}{\sqrt{2\pi}} e^{inx} \frac{1}{\sqrt{2\pi}} e^{imx} \, dx \\
&= \frac{1}{2\pi} \int_{\bar{T}} e^{i(m-n)x} \, dx \\
&= \begin{cases} \frac{1}{2\pi} \int_{\bar{T}} 1 \, dx = 1 & m = n \\ \frac{1}{2\pi} \frac{1}{i(m-n)} e^{i(m-n)x}|_0^{2\pi} = 0 & m \neq n \end{cases} \\
&= \delta_{mn}(x) \implies \text{ orthonormal}
\end{aligned}$$
- Completeness of $\{e_n\}$.
That is, *span*$\{e_n : n \in \mathbb{Z}\}$ is dense in $L^2(\bar{T})$. Any finite linear combination of (e_n) is called a trigonometric polynomial. So, we need to show: Trigonometric polynomials are dense in $L^2(\bar{T})$ with respect to $\|\cdot\|_2$.

  $\Leftarrow$ we only really need to show Trigonometric are dense in $C(\bar{T})$ with respect to $\|\cdot\|_2$ because $C(\bar{T})$ is dense in $L^2(\bar{T})$.

$\Leftarrow$ we only really need to show Trigonometric are dense in $C(\bar{T})$ with respect to $\|\cdot\|_\infty$ because

$$\|f - f_n\|_2 = \left(\int_{\bar{T}} |f - f_n|^2 \, dx\right)^{\frac{1}{2}} \leq \sqrt{2\pi}\, \|f - f_n\|_\infty$$

$\Leftarrow$ find ϕ_n such that $f_n = \phi_n * f$ is a trigonometric polynomial and $f_n \to f$ uniformly.

Definition 10.2. *A family of functions $\{\phi_n \in C(\bar{T})\}_{n\in\mathbb{N}}$ is called an approximate identity if*

1. $\phi_n(x) \geq 0$
2. $\int_{\bar{T}} \phi_n(x)\, dx = 1$ *for all* n.
3. $\forall 0 < \delta < \pi$, $\lim_{n\to\infty} \int_{\delta\leq|x|\leq\pi} \phi_n(x)\, dx = 0$

Note 10.3. *The convolution*

$$(\phi_n * f)(x) = \int_{\bar{T}} \underbrace{f(y)\phi_n(x-y)}_{\text{weighted average of } f} \quad dy$$

And this puts more weights on values near x. Therefore, as $n \to \infty$, we hope

$$(\phi_n * f)(x) \to x$$

Theorem 10.4. *If $\{\phi_n \in C(\bar{T})\}$ is an approximate identity and $f \in C(\bar{T})$, then*

$$(\phi_n * f)(x) \to f$$

uniformly as $x \to \infty$.

Proof. For all $\epsilon > 0$, since f is continuous on $[-\pi, \pi]$ it is also uniformly continuous. Thus, $\exists\delta > 0$ such that

$$|f(x) - f(y) < \epsilon \text{ whenever } |x - y| < \delta$$

Now, for all $x \in \bar{T}$,

$$\begin{aligned}
|(\phi_n * f)(x) - f(x)| &= \left|\int_{-\pi}^{\pi} \phi_n(y) f(x-y)\, dy - f(x)\int_{-\pi}^{\pi} \phi_n(x)\, dy\right| \\
&= \left|\int_{-\pi}^{\pi} \phi_n(y)\left(f(x-y) - f(x)\right)\, dy\right| \\
&\leq \int_{|y|<\delta} \phi_n(y)\left|f(x-y) - f(x)\right|\, dy + \leq \int_{|y|>\delta} \phi_n(y)\left|f(x-y) - f(x)\right|\, dy \\
&\leq \epsilon \int_{|y|<\delta} \phi_n(y)\, dy + 2M \int_{|y|\geq\delta} \phi_n(y)\, dy
\end{aligned}$$

where M is an upper bound of $|f|$ on $\bar{T}$. Therefore,

$$\lim_{n\to\infty} \|\phi_n * f - f\|_\infty \leq \epsilon + 2M \cdot 0 = \epsilon$$

Since ϵ is arbitrary, then $\|\phi_n * f - f\|_\infty \to \epsilon$. ∎

Theorem 10.5. *The trigonometric polynomials are dense in $C(\bar{T})$ with respect to $\|\cdot\|_\infty$.*

Proof. For all $n \in \mathbb{N}$, let

$$\phi_n(x) = c_n(1 + \cos x)^n \text{ where } c_n \text{ is chosen such that } \int_{\bar{T}} \phi_n(x)\, dx = 1$$

Some properties of this function are:

1. $\phi_n \geq 0$
2. $\int_{\bar{T}} \phi_n(x)\ dx = 1$
3. $\forall \delta$, **Homework problem**

So $\{\phi_n\}$ is an approximate identity. Thus, by the previous theorem, we know:

$$\phi_n * f \to f \text{ uniformly on } \bar{T}$$

Notice, each ϕ_n is itself a trigonometric polynomial:

$$\begin{aligned}
\phi_n(x) = c_n(1+\cos x)^n &= c_n\left(1 + \frac{e^{ix}+e^{-ix}}{2}\right) \\
&= \frac{c_n}{2^n}(e^{i\frac{x}{2}} + e^{-i\frac{x}{2}})^n \\
&= \frac{c_n}{2^n}\sum_{k=0}^{2n}\binom{2n}{n}(e^{i\frac{x}{2}})^k(e^{-i\frac{x}{2}})^{2n-k} \\
&= \frac{c_n}{2^n}\sum_{k=0}^{2n}\binom{2n}{n}e^{i(k-n)x} \\
&= \frac{c_n}{2^n}\sum_{k=-n}^{n}\binom{2n}{n+k}e^{ikx} \\
&= \sum_{k=-n}^{n} a_{kn}e^{ikx}
\end{aligned}$$

Second, we see:

$$\begin{aligned}
(\phi_n * f)(x) &= \int_{\bar{T}} \phi_n(x-y)f(y)\ dy \\
&= \int_{\bar{T}} \sum_{k=-n}^{n} a_{kn}e^{ik(x-y)}f(y)\ dy \\
&= \sum_{k=-n}^{n} a_{kn}e^{ikx}\int_{\bar{T}} e^{-iky)}f(y)\ dy \\
&= \sum_{k=-n}^{n} b_{kn}e^{ikx}
\end{aligned}$$

is also a trigonometric polynomial. Therefore, the trigonometric polynomials are dense in $C(\bar{T})$! And because of uniformity, this extends into L^2. ■

Corollary 10.5.1. $\{e_n\}_{n=1}^{\infty}$ *forms an orthonormal basis for* $L^2(\bar{T})$

As a result, we see that for every function $f \in L^2(\bar{T})$ can be represented as:

$$f(x) = \sum_{n=-\infty}^{\infty}\langle e_n, f\rangle e_n = \frac{1}{\sqrt{2\pi}}\sum_{n=-\infty}^{\infty}\hat{f}e^{inx}$$

where $\hat{f} = \langle e_n, f\rangle = \frac{1}{\sqrt{2\pi}}\int_{\bar{T}} f(x)e^{-inx}\ dx \in \mathbb{C}$ are called the Fourier Coefficients of f. Here, the equality means the convergence of the partial sums to f in the L^2 norm:

$$\lim_{N\to\infty}\int_{\bar{T}}\left|\left(\frac{1}{\sqrt{2\pi}}\sum_{n=-N}^{N}\hat{f}e^{inx}\right) - f(x)\right|^2\ dx = 0$$

Moreover, by Parseval's identity holds:

$$\|f\|_{L^2}^2 = \sum_{n=-\infty}^{\infty} |\hat{f}_n|^2$$

As well as the Generalized Parseval's Identity:

$$\langle f, g\rangle = \sum_{n=-\infty}^{\infty} \overline{\hat{f}_n} \cdot \hat{g}_n$$

The periodic Fourier transform $\mathcal{F} : L^2(\bar{T}) \to \ell^2(\mathbb{Z})$

$$f \to \hat{f}_n$$

is a Hilbert space isomorphism.

Theorem 10.6. *If $f, g \in L^2(\bar{T})$, then*

$$\widehat{(f * g)}_n = \sqrt{2\pi}\hat{f}_n \cdot \hat{g}_n$$

Proof. Since $L^2(\bar{T}) \subset L^1(\bar{T})$, both f and g are integrable. So:

$$\begin{aligned}
\widehat{(f * g)}_n &= \frac{1}{\sqrt{2\pi}} \int_{\bar{T}} (f * g)(x) e^{-inx}\, dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{\bar{T}} \left(\int_{\bar{T}} f(x-y) g(y)\, dy \right) e^{-inx}\, dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{\bar{T}} \left(\int_{\bar{T}} f(x-y) g(y) e^{-inx}\, dx \right) dy && \text{Fubini} \\
&= \frac{1}{\sqrt{2\pi}} \int_{\bar{T}} \left(\int_{\bar{T}} f(x-y) e^{-in(x-y)}\, dx \right) g(y) e^{-iny}\, dy \\
&= \hat{f}_n \int_{\bar{T}} g(y) e^{-iny}\, dy = \sqrt{2\pi}\hat{f}_n \hat{g}_n
\end{aligned}$$

■

Remark 10.7. *Sometimes, it maybe not easy to calculate $f * g$ directly. Instead, follow the steps:*

1. *Calculate $\hat{f}_n, \hat{g}_n$ for each n*
2. $\widehat{(f * g)}_n = \sqrt{2\pi}\hat{f}_n \cdot \hat{g}_n$
3. $f * g(x) = \sum_{n=-\infty}^{\infty} \widehat{(f * g)}_n \frac{1}{\sqrt{2\pi}} e^{inx}$

Question 10.8. *Given $f \in L^2(\bar{T})$, what is the best approximation of f by a trigonometric polynomial of degree N in $\|\cdot\|_2$-norm?*

Answer:

$$f_N(x) = \frac{1}{\sqrt{2\pi}} \sum_{n=-N}^{N} \hat{f}_n e^{inx}$$

Why? Let's recall orthogonal projections, with $\mathcal{H}$:= a Hilbert space, and M := a closed linear subspace. For every $x \in \mathcal{H}, \exists$ a unique $x_m \in M$ such that

$$\|x - x_m\| = \inf_{y \in M} \|x - y\|$$

Here, $x - x_m \in M^{\perp}$. For the case of our current approximation, $\mathcal{H} = L^2(\bar{T})$ and $M := \{$ trigonometric polynomials of deg
Then, for all $f = \sum_{n=-\infty}^{\infty} \hat{f}_n e_n \in L^2(\bar{T})$, be see

- $f_M \in M \implies f_M$ does not contain multiples of e_n where $|n| > N$
- $f - f_M \in M^\perp \implies f - f_M$ does not contain e_n terms with $|n| \le N$

Therefore,

$$f_M = \sum_{n=-N}^{N} \hat{f}_n e_n$$

To summarize, in the space $L^2(\bar{T}, \mathbb{C})$ we see:

1. Has an orthonormal basis
$$e_n = \frac{1}{\sqrt{2\pi}} e^{inx}$$

2. Every function $f \in L^2(\bar{T})$ can be respresented
$$f = \sum_{n=-\infty}^{\infty} \hat{f}_n e_n$$
where $\hat{f}_n = \frac{1}{\sqrt{2\pi}} \int_{\bar{T}} f(x) e^{-inx}\ dx$

Now, for any $f \in L^2(\bar{T}, \mathbb{R}) \subset L^2(\bar{T}, \mathbb{C})$ we see that

$$f = \sum_{n=-\infty}^{\infty} \hat{f}_n e_n$$

But also,

$$\begin{aligned}
f = Re[f] &= \sum_{n=-\infty}^{\infty} \Big(Re[\hat{f}_n] Re[e_n] - Im[\hat{f}_n] Im[e_n] \Big) \\
&= \sum_{n=-\infty}^{\infty} \left[\frac{1}{\sqrt{2\pi}} \left(\int_{\bar{T}} f(x) \cos nx\ dx \right) \frac{1}{\sqrt{2\pi}} \cos nx + \frac{1}{\sqrt{2\pi}} \left(\int_{\bar{T}} f(x) \sin nx\ dx \right) \frac{1}{\sqrt{2\pi}} \sin nx \right] \\
&= \frac{1}{2\pi} \int_{\bar{T}} f(x)\ dx + \sum_{n=1}^{\infty} \frac{1}{\pi} \left[\int_{\bar{T}} f(x) \cos nx\ dx \cos(nx) + \int_{\bar{T}} f(x) \sin nx\ dx \sin(nx) \right] \\
&= \frac{1}{2\pi} a_0 + \frac{1}{\pi} \sum_{n=1}^{\infty} a_n \cos nx + b_n \sin nx
\end{aligned}$$

11 Sobolev Spaces

Recall the Fourier transform:

$$L^2(\mathbb{T}) \to \ell^2(\mathbb{Z})$$
$$f \to (\hat{f}_n)$$

is a Hilbert space isomorphism. Many calculations on the right side is much easier.

Example 11.1 (Convolution)**.**

$$f * g = \int_{\bar{\mathbb{T}}} f(x-y)g(y)\ dy \quad \widehat{(f*g)} = \sqrt{2\pi}\hat{f}_n \cdot \hat{g}_n$$

Lemma 11.2. *Suppose that $f \in C^1(\bar{\mathbb{T}})$, then*

$$\widehat{f'}_n = in\hat{f}_n$$

Moreover, if $f \in C^k(\bar{\mathbb{T}})$, then $\widehat{f^{(k)}}_n = (in)^k\hat{f}_n$.

Proof.

$$\begin{aligned} \hat{f'}_n &= \frac{1}{\sqrt{2\pi}}\int_{\bar{T}} f'(x)e^{-inx}\ dx \\ &= \left[\frac{1}{\sqrt{2\pi}}f(x)e^{-inx}\right]_{-\pi}^{\pi} - \frac{-in}{\sqrt{2\pi}}\int_{\bar{T}} f(x)e^{-inx}\ dx \\ &= (in)\hat{f}_n \end{aligned}$$

■

Note 11.3. *Notice that we typically need limits to calculate the derivative of f. Instead, to calculate $\widehat{f'}_n$, we only need $in\hat{f}_n \in \ell^2(\mathbb{Z})$. This condition is weaker, leading us to the concept of a weak derivative of f. Specifically, take any $f \in L^2(\bar{T})$, map this f to $\hat{f}_n$. If $(in\hat{f})n) \in \ell(\mathbb{Z})$, then return to $L^2(\bar{T})$ gives us a* weak derivative*.*

Definition 11.4. *The Sobolev Space is defined*

$$H^1(\bar{T}) = \{f \in L^2(\bar{T}) : \sum_{n=-\infty}^{\infty} n^2|\hat{f}_n|^2 < +\infty\}$$

Definition 11.5. *For any $f \in H^1(\bar{T})$, the weak L^2-derivative $f' \in L^2(\bar{T})$ is defined by the L^2-convergent Fourier series*

$$f'(x) = \frac{1}{\sqrt{2\pi}}\sum_{n=-\infty}^{\infty}(in)\hat{f}_n e^{inx}$$

Notice, $H^1(\bar{T})$ is a Hilbert space, with respect to the inner product defined:

$$\begin{aligned} \langle f, g\rangle_{H^1} &:= \langle f, g\rangle_{L^2} + \langle f', g'\rangle_{L^2} \\ &= \int_{\bar{T}} \overline{f}\cdot g\ dx + \int_{\bar{T}} \overline{f}'\cdot g'\ dx \\ &= \sum_{n=-\infty}^{\infty}\overline{\hat{f}_n}\hat{g}_n + \sum_{n=-\infty}^{\infty}\overline{in\hat{f}_n}(in\hat{g}_n) && \text{Parseval's Identity} \\ &= \sum_{n=-\infty}^{\infty}(1+n^2)\hat{f}_n\hat{g}_n \end{aligned}$$

In particular,

$$\|f\|_{H^1}^2 = \|f\|_{L^2}^2 + \|f'\|_{L^2}^2 = \sum_{n=-\infty}^{\infty} (1+n^2)|\hat{f}_n|^2$$

Exercise 11.6. *Check $H^1(\bar{T})$ is a Hilbert space and $C^1(\bar{T})$ is dense in $H^1(\bar{T})$ under the norm $\|\cdot\|_{H^1}$*

11.1 Integration by Parts

For any $f, g \in H^1(\bar{T})$,

$$\langle f', g\rangle = \sum_{n=-\infty}^{\infty} \overline{in\hat{f}_n}\hat{g}_n = -\sum_{n=-\infty}^{\infty} \overline{\hat{f}_n}(in\hat{g}_n) = -\langle f, g'\rangle$$

Therefore,

$$\int_{\bar{T}} f'g \ dx = \langle f', g\rangle = -\langle f, g'\rangle = \int_{\bar{T}} fg' \ dx$$

In particular, for any $\phi \in C^1(\bar{T})$,

$$\int_{\bar{T}} f'\phi \ dx = -\int_{\bar{T}} f\phi' \ dx$$

Remark 11.7. *There is an equivalent way to define weak L^2 derivatives.*

Definition 11.8. *A function $f \in L^2(\bar{T})$ belongs to $H^1(\bar{T})$ if*

$$\left|f\phi' \ dx\right| \leq M \|\phi\|_{L^2} \quad \forall \phi \in C^1(\bar{T})$$

That is,

$$F(\phi) := \int_{\bar{T}} f\phi' \ dx$$

is a bounded linear operator on $C^1(\bar{T})$. This implies $F \in (L^2(\bar{T}))^* = L^2(\bar{T}) \implies \exists g \in L^2(\bar{T})$ such that $F(\phi) = \langle g, \phi\rangle \implies \int_{\bar{T}} f\phi' \ dx = \int_{\bar{T}} \bar{g}\phi \ dx$. If $f \in H^1(\bar{T})$, then the weak derivative of f is the unique element of $L^2(\bar{T})$ such that

$$\int_{\bar{T}} f'\phi \ dx = \int_{\bar{T}} f\phi' \ dx$$

for all test functions $\phi \in C^1(\bar{T})$.

11.2 Smoothness of Functions

Notice, when $f \in L^2(\bar{T})$, we see:

$$(\hat{f}_n \in \ell^2(\mathbb{Z}) \implies \sum_{n=-\infty}^{\infty} |\hat{f}_n|^2 < \infty \implies |\hat{f}_n| \to 0 \text{ as } n \to \infty \implies |\hat{f}_n| \in O(1)$$

On the other hand, when $f \in H^1(\bar{T})$, we see:

$$(in\hat{f}_n \in \ell^2(\mathbb{Z}) \implies |n||\hat{f}_n| \to 0 \implies |\hat{f}_n| \in O(\frac{1}{|n|})$$

In general, smoothness of f can be identified with the rate of decay of the Fourier coefficients of $(\hat{f}_n)$. Therefore, the smoother f is, the faster these coefficients decay and vice versa. Notice:

$$f \in C^k(\bar{T}) \implies \hat{f_n^{(k)}} = (in)^k\hat{f}_n \in \ell^2(\mathbb{Z}) \implies |\hat{f}_n| \in O(\frac{1}{|n|^k})$$

Definition 11.9. *For any $k = 1, 2, \ldots,$ we define the general Sobolev Space*

$$H^k(\bar{T}) = \{f \in L^2(\bar{T}) : \sum_{n=-\infty}^{\infty} |n|^{2k} |\hat{f}_n|^2 < +\infty\}$$

Note 11.10. *If $f \in H^k(\bar{T}) \implies f' \in H^{k-1}(\bar{T}) \implies \ldots \implies f^{(k-1)} \in H^1(\bar{T})$.*

In general, for any $s \geq 0$, we define the Sobolev space

$$H^s(\bar{T}) = \{f \in L^2(\bar{T}) : \sum_{n=-\infty}^{\infty} |n|^{2s} |\hat{f}_n|^2 + \infty\}$$

For any $f \in L^2(\bar{T})$, we may define an operator of fractional order differentiation of order s by

$$f^{(s)}(x) := \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} (in)^s \hat{f}_n e^{inx}$$

Thus,

$$H^s(\bar{T}) = \{f \in L^2(\bar{T}) : \left\| f^{(s)} \right\|_{L^2} < \infty\}$$

Lemma 11.11. *Suppose that $f \in H^s(\mathbb{T})$ for $s > \frac{1}{2}$. Let*

$$S_N(x) = \frac{1}{\sqrt{2\pi}} \sum_{n=-N}^{N} \hat{f}_n e^{inx}$$

be the Nth partial sum of the Fourier series of f. Then

$$\|S_N - f\|_\infty \leq \frac{C_s}{N^{s-\frac{1}{2}}} \left\| f^{(s)} \right\|_{L^2}$$

and S_N converges to f uniformly. In particular, f is continuous.

Proof. For any $M \geq N$,

$$\begin{aligned}
\|S_N - S_M\|_\infty &= \left\| \frac{1}{\sqrt{2\pi}} \sum_{N<|n|\leq M} \hat{f}_n e^{inx} \right\| \\
&\leq \frac{1}{\sqrt{2\pi}} \sum_{N<|n|\leq M} |\hat{f}_n| \\
&= \frac{1}{\sqrt{2\pi}} \sum_{N<|n|\leq M} |n|^s |\hat{f}_n| \frac{1}{|n|^s} \\
&\leq \frac{1}{\sqrt{2\pi}} \left[\sum_{N<|n|\leq M} |n|^{2s} |\hat{f}_n|^2 \right]^{\frac{1}{2}} \left[\sum_{N<|n|\leq M} \frac{1}{|n|^{2s}} \right]^{\frac{1}{2}} && \text{Cauchy Schwarz} \\
&\leq \frac{1}{\sqrt{2\pi}} \left\| f^{(s)} \right\|_{L^2} \left[2 \int_N^\infty \frac{1}{r^{2s}} \, dr \right]^{\frac{1}{2}} \\
&= \frac{1}{\sqrt{\pi}} \left\| f^{(s)} \right\|_{L^2} \left[\frac{r^{1-2s}}{1-2s} \Big|_N^\infty \right]^{\frac{1}{2}} \\
&= \frac{1}{\sqrt{\pi}} \left\| f^{(s)} \right\|_{L^2} \sqrt{\frac{N^{1-2s}}{2s-1}} \\
&= \frac{C_s}{N^{s-\frac{1}{2}}} \left\| f^{(s)} \right\|_{L^2}
\end{aligned}$$

Therefore, $\{S_N\}$ is Cauchy in $(C(\mathbb{T}), \|\cdot\|_\infty)$. Let $M \to \infty$ shows that the set of trigonometric functions is dense! ∎

11.3 Sobolev Embeddings

If $f \in H^s(\bar{T})$ for $s > k + \frac{1}{2}, k \in \mathbb{N} \cup \{0\}$, then $f \in C^k(\bar{T})$. This can be seen since:

$$f \in H^s(\bar{T}) \implies f' \in H^{s-1}(\bar{T}) \implies f'' \in H^{s-2}(\bar{T}) \implies \dots \implies f^{(k)} \in H^{s-k > \frac{1}{2}}(\bar{T})$$

In general, if $f \in H^s(\bar{T})$, then

- $f \in C(\bar{T}^d)$ when $s > \frac{d}{2}$
- $f \in C^k(\bar{T}^d)$ when $s > k + \frac{d}{2}$

Roughly speaking, there is a "loss" of slightly more than $\frac{1}{2}$ a derivative per space dimension in passing from L^2 derivative to continuous derivative.

11.4 Applications of Fourier Series

11.4.1 The Laplace Equation

Solving Dirichlet Problem

$$\begin{cases} \Delta u = 0 & \text{in unit disk } D \subset \mathbb{R}^2 \\ u = f & \text{on } \partial D \end{cases}$$

$\Delta u = 0$ is called the Laplace equation, where

$$\Delta u = u_{xx} + u_{yy} = u_{rr} + \frac{1}{r} u_r + \frac{1}{r^2} u_{\theta\theta}$$

A solution of $\Delta u = 0$ is called a harmonic function.
The PDE is linear since for $i = 1, 2$

$$\begin{cases} \Delta u_i = 0 & \text{in unit disk } D \subset \mathbb{R}^2 \\ u_i = f_i & \text{on } \partial D \end{cases} \implies \begin{cases} \Delta (u_1 + u_2) = 0 & \text{in unit disk } D \subset \mathbb{R}^2 \\ u_1 + u_2 = f_1 + f_2 & \text{on } \partial D \end{cases}$$

To solve the Dirichlet problem for some boundary condition f, then

- $f = \text{const} \implies u = c$
- $f = \cos\theta \implies u = r\cos\theta = x$
- $f = \sin\theta \implies u = r\sin\theta = y$
- $f = e^{in\theta} \implies u = r^{|n|} e^{in\theta}$
- $f = \sum_{n=-\infty}^{\infty} a_n e^{in\theta}$ where

  $$a_n = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-ins} \, dx$$

  then

  $$u(r,\theta) = \sum_{n=-\infty}^{\infty} a_n r^{|n|} e^{in\theta}$$

  In particular, if f is real-valued

  $$u(r,\theta) = \frac{a_0}{2} + \sum_{n=1}^{\infty} r^n (a_n \cos n\theta + b_n \sin n\theta)$$

Now comes an amazing fact: The series can be summed explicitly!

$$\begin{aligned}
u(r,\theta) &= \sum_{n=-\infty}^{\infty} a_n r^{|n|} e^{in\theta} \\
&= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \left(\int_0^{2\pi} e^{-ins} f(s)\ ds \right) r^{|n|} e^{in\theta} \\
&= \frac{1}{2\pi} \int_0^{2\pi} \left(\sum_{n=-\infty}^{\infty} r^{|n|} e^{in(\theta-s)} f(s)\ ds \right) \\
&= \frac{1}{2\pi} P_r * f
\end{aligned}$$

where

$$\begin{aligned}
P_r(\theta) &= \sum_{n=-\infty}^{\infty} r^| n| e^{in\theta} \\
&= \sum_{n=-0}^{\infty} r^{|n|} e^{in\theta} + \sum_{n=1}^{\infty} r^{|n|} e^{-in\theta} \\
&= \frac{1}{1-re^{i\theta}} + \frac{re^{-i\theta}}{1-re^{-i\theta}} \\
&= \frac{1-r^2}{1-2r\cos\theta + r^2}
\end{aligned}$$

This is called the Poisson kernel. Here, we have the Poisson's formula:

$$u(r,\theta) = \frac{1}{2\pi} P_r * f = \frac{1}{2\pi} \int_0^{2\pi} \frac{1-r^2}{1-2r\cos(\theta-s)+r^2} f(s)\ ds$$

In can be written in a more geometric way. By the law of cosine:

$$|x-x'|^2 = 1 + r^2 - 2r\cos(\theta - s)$$

So

$$u(x) = \frac{1}{2\pi} \int_0^{2\pi} \frac{1-|x|^2}{|x-x'|^2} f(s)\ ds = \frac{1-|x|^2}{2\pi} \int_{|x|=1} \frac{u(x')}{|x-x'|^2} ds$$

In general, for $D = \{r < a\}$

$$\begin{aligned}
u(x) &= \frac{a^2-|x|^2}{2\pi a} \int_{|x'|=a} \frac{u(x')}{|x-x'|^2}\ ds \\
&= \frac{1}{2\pi} \int_0^{2\pi} \frac{(a^2-r^2) f(s)}{a^2 - 2ar\cos(\theta-s) + r^2}\ ds
\end{aligned}$$

In particular,

$$u(0) = \frac{a^2}{2\pi a} \int_{|x'|=a} \frac{u(x')}{a^2}\ ds = \int_{2\pi a} \int_{|x'|=a} u(x')\ ds$$

Mean-Value Property: Let u be a harmonic function on a disk D. Then the value of u at the center of D equals to the average of u on its circumference.

11.4.2 The Heat Equation

We can solve the heat equation

$$\begin{cases} u_t = u_{xx} \\ u(x,0) = f(x) \end{cases}$$

In it's conception, $u(x,t)$ represented the temperature at location x at time t. To solve the Dirichlet problem for example boundary conditions f, then

- $f = e^{inx} \implies$ we can make a guess of $u(x,t) = g(t)e^{inx}$

  $$\implies u_t = g'(t)e^{inx}$$

  $$\implies u_{xx} = g(t)(in)^2 e^{inx}$$

  Therefore,

  $$g'(t) = g(t)(in)^2 = -n^2 g(t) \implies g(t) = e^{-n^2 t}$$

  That is,

  $$u(x,t) = e^{-n^2 t} e^{inx}$$

- $f(x) = \sum_{n=-N}^{N} a_n e^{inx}$ Then by linearity

  $$u(x,t) = \sum_{n=-N}^{N} a_n e^{-n^2 t} e^{inx}$$

  Note 11.12.

  $$\|u(x,t)\|_{L^2(\mathbb{T})}^2 = \left\| \sum_{n=-N}^{N} a_n e^{-n^2 t} e^{inx} \right\|_{L^2(\mathbb{T})}^2 = \sum_{n=-N}^{N} |a_n|^2 e^{-2n^2 t} 2\pi \le \|u(x,t=0)\|^2$$

  So the norm decreases as time goes on.

- $f(x) \in L^2(\mathbb{T})$. Then

  $$f(x) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} \hat{f}_n e^{inx}$$

  Thus

  $$\begin{aligned} u(x,t) &= \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} \hat{f}_n e^{-n^2 t} e^{inx} \\ &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \int_{\mathbb{T}} f(s) e^{-ins}\, ds e^{-n^2 t} \\ &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \int_{\mathbb{T}} f(s) e^{in(x-s)}\, dx e^{-n^2 t} \\ &= g * f \end{aligned}$$

  where

  $$g = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{inx} e^{-n^2 t}$$

Claim: When $t > 0, u(x,t) \in H^k(\mathbb{T})$ for all k. Indeed

$$\begin{aligned}
\sum_{n=-\infty}^{\infty} |n|^{2k} |\widehat{u(x,t)}_n|^2 &= \sum_{n=-\infty}^{\infty} |n|^{2k} \left(|\hat{f}_n| e^{-n^2 t} \right)^2 \\
&= \sum_{n=-\infty}^{\infty} \frac{|n|^{2k}}{e^{2n^2 t}} |\hat{f}_n|^2 \\
& \max_{|n|<\infty} \left\{ \frac{|n|^{2k}}{e^{2n^2 t}} \right\} \|f\|_{L^2}^2 < \infty
\end{aligned}$$

By Sobolev embedding, $u(\cdot, t) \in C^\infty(\mathbb{T})$ for all $t > 0$.
Now, let

$$\langle u \rangle = \frac{1}{2\pi} \int_{\mathbb{T}} f(x)\ dx$$

by the mean temperature over $\mathbb{T}$.

Note 11.13.

$$\frac{d}{dt} \int_{\mathbb{T}} u(x,t)\ dx = \int_{\mathbb{T}} \frac{\partial}{\partial t} u(x,t)\ dx = \int_{\mathbb{T}} u_t(x,t)\ dx = \int_{\mathbb{T}} u_{xx}(x,t)\ dx = 0$$

Therefore, the function is constant with respect to t and

$$\int_{\mathbb{T}} u(x,t)\ dx = \int_{\mathbb{T}} u(x,0) dx = \langle u \rangle$$

Claim: $u(x,t) \to \langle u \rangle$ exponentially fast as $t \to \infty$

$$\begin{aligned}
\|u(x,t) - \langle u \rangle\|_\infty &= \left\| \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} \hat{f}_n e^{-n^2 t} e^{inx} - f_0 \right\|_\infty \\
&= \left\| \frac{1}{\sqrt{2\pi}} \sum_{n \neq 0} \hat{f}_n e^{-n^2 t} e^{inx} \right\|_\infty \\
&\leq \frac{1}{\sqrt{2\pi}} \sum_{n \neq 0} |\hat{f}_n e^{-n^2 t}| \\
&\leq \sqrt{\frac{1}{\sqrt{2\pi}} \sum_{n \neq 0} |\hat{f}_n|^2} \sqrt{\sum_{n \neq 0} |e^{-n^2 t}|^2} && \text{Cauchy Schwarz} \\
&= \sqrt{\frac{1}{\sqrt{2\pi}} \sum_{n \neq 0} |\hat{f}_n|^2} \sqrt{2 \sum_{n=1}^{\infty} |e^{-n^2 t}|^2} \\
&\leq C \|f\|_{L^2} \sqrt{\frac{e^{-2t}}{1 - e^{-2t}}} \\
&\leq C_2 \|f\|_{L^2} e^{-t}
\end{aligned}$$

11.4.3 Isoperimetric Inequality

Suppose that Γ is a simple, closed curve in $\mathbb{R}^2$ of length ℓ and let A denote the area of the region enclosed by this curve Γ. Then

$$4\pi A \leq \ell^2$$

with equality if and only if Γ is a circle.

Note 11.14. *For a circle* $\ell = 2\pi r, A = \pi r^2$ *so* $2\pi A = 4\pi - \pi r^2 = (2\pi r)^2 = \ell^2$

Proof. Without loss of generality, we may assume that $\ell = 2\pi$. We need to show

$$4\pi A \leq (2\pi)^2 = 4\pi^2 \iff A \leq \pi$$

Let $\gamma : [0, 2\pi] \to \mathbb{R}^2$ with

$$\gamma(s) = (x(s), y(s))$$

by a parametrization by arc-length of Γ. That is,

$$(x')^2 + (y')^2 = 1 \implies \frac{1}{2\pi} \int_0^{2\pi} (x'(s))^2 + (y'(s))^2 \, ds = 1$$

∎

Consider the Fourier expansions of $x(s)$ and $y(s)$.

$$\begin{cases} x(s) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} a_n e^{ins} \\ y(s) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} b_n e^{ins} \end{cases}$$

Then

$$\begin{cases} x'(s) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} (in) a_n e^{ins} \\ y'(s) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} (in) b_n e^{ins} \end{cases}$$

By Parseval's Identity, we see

$$\sum_{n=-\infty}^{\infty} n^2 (|a_n|^2 + |b_n|^2) = 2\pi$$

By Green's Formula,

$$\int_{\partial\Omega} u \; dy - v \; dx = \int_{\Omega} (u_x + v_y) dx dy$$

$$\implies \int_{\mathbb{T}} (x(s)y'(s) - y(s)x'(s)) \; ds = \int_A (1+1) dx dy = 2A$$

So

$$\begin{aligned} A &= \frac{1}{2} \int_{\mathbb{T}} (x(s)y'(s) - y(s)x'(s)) \; ds \\ &= \frac{1}{2} \sum_{n=-\infty}^{\infty} (a_n (in) bn - b_n (in) a_n) \\ &\leq \sum_{n=-\infty}^{\infty} |n||a_n||b_n| \\ &\leq \sum_{n=-\infty}^{\infty} n^2 \frac{|a_n|^2 + |b_n|^2}{2} = \pi \end{aligned}$$

Lastly, when $A = \pi$, we have $|n| = n^2 \implies n = \pm 1 \implies x(s) = \frac{1}{\sqrt{2\pi}} (a_{-1} e^{-is} + a_0 + a_1 e^{is}) = a_0 + \cos(\alpha + s)$

$$\implies y(s) = b_0 + \sin(\alpha + s)$$

This is the parametrization of a circle!

12 Bounded Linear Operators on a Hilbert Space

12.1 The Dual of a Hilbert Space

Recall, if we let X be a linear space, then the dual space $X^* = B(X, \mathbb{C}) = \{$ bounded linear functionals on $X\}$ is a Banach space. If X is also a Banach space, then $X \subset X^{**}$. When $X = X^{**}$, then X is reflexive.

Now, if $\mathcal{H}$ is a Hilbert space, we'll show that $\mathcal{H}^* = \mathcal{H}$.
Let $\mathcal{H}$ be a Hilbert space. For all $y \in \mathcal{H}$, define $\phi_y : \mathcal{H} \to \mathbb{C}$ such that

1. ϕ_y is linear
2. $|\phi_y(x)| = |\langle y, x\rangle| \leq \|y\| \, \|x\| \implies \|\phi_y(y)\| = \sup_{x\neq 0} \frac{|\phi_y(x)|}{\|x\|} \leq \|y\|$ On the other hand,
$$|\phi_y(y)| = |\langle y, y\rangle| = \|y\|^2 \implies \|\phi_y\| = \|y\|$$

So, we get a map

$$J : \mathcal{H} \to \mathcal{H}^*$$
$$y \to \phi_y$$

Notice, J is one-to-one, because

$$\phi_y = 0 \implies \langle y, x\rangle = 0 \forall x \implies \langle y, y\rangle = 0 \implies y = 0$$

Now, we'll show that J is onto. That is, $\forall \phi \in \mathcal{H}^*, \exists y \in \mathcal{H}$ such that $\phi = \phi_y$.

Theorem 12.1 (Riesz Representation Theorem)**.** *If ϕ is a bounded linear functional on a Hilbert space $\mathcal{H}$ (i.e. $\phi \in \mathcal{H}^*$) then there exists a unique vector y in$\mathcal{H}$ such that*

$$\phi(x) = \langle y, x\rangle$$

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

<u>Aside:</u> If such a y exists, then for all $x \in \mathrm{Ker}(\phi)$,

$$0 = \phi(x) = \langle y, x\rangle \implies y \perp \mathrm{Ker}(\phi) \implies y \in (\mathrm{Ker}(\phi))^\perp$$

Proof. (Existence) If $\phi = 0$, then $y = 0$. So we may assume $\phi \neq 0$. This this case $\mathrm{Ker}(\phi)$ is a proper, closed subspace of $\mathcal{H}$. Thus, we can decompose $\mathcal{H}$ as

$$\mathcal{H} = \mathrm{Ker}(\phi) \oplus (\mathrm{Ker}(\phi))^\perp$$

as a direct sum. Now, pick any $z \in (\mathrm{Ker}(\phi))^\perp$ with $z \neq 0$. Then for all $x \in \mathcal{H}$,

$$x = (x - \frac{\phi(x)z}{\phi(z)}z) + \frac{\phi(x)z}{\phi(z)}z \in \mathrm{Ker}(\phi) \oplus (\mathrm{Ker}(\phi))^\perp$$

$$\implies \langle z, x\rangle = \langle z, x - \frac{\phi(x)}{\phi(z)}z\rangle + \langle z, \frac{\phi(x)}{\phi(z)}z\rangle = 0 + \frac{\phi(x)}{\phi(z)}\|z\|^2$$

$$\implies \phi(x) = \frac{\phi(z)}{\|z\|^2}\langle z, y\rangle = \langle \overline{\frac{\phi(z)}{\|z\|^2}}z, x\rangle = \langle y, x\rangle$$

where $y = \overline{\frac{\phi(z)}{\|z\|^2}}z \implies \phi = \phi_y$.
(Uniqueness) Suppose $\phi = \phi_y = \phi_{\tilde{y}}$ for some $y, \tilde{y} \in \mathcal{H}$. Then

$$\langle y - \tilde{y}, x\rangle = \langle y, x\rangle - \langle \tilde{y}, x\rangle = \phi_y(x) - \phi_{\tilde{y}}(x) = 0$$

In particular,

$$\langle y - \tilde{y}, y - \tilde{y}\rangle = 0 \implies y = \tilde{y}$$

∎

Example 12.2. $\mathcal{H} = \mathbb{R}^3, \phi \in \mathcal{H}^*$. *Then any linear bounded function can be written:*

$$\phi(x, , y, z) = ax + by + cz = (a, b, c) \cdot (x, y, z)$$

Example 12.3. $\mathcal{H} = L^2(\mathbb{T})$. *For any $n \in \mathbb{Z}$, then any linear bounded function $\phi_n : L^2(\mathbb{T}) \to \mathbb{C}$ can be written:*

$$f \to_{\phi_n} \hat{f}_n = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} f(x) e^{-inx}\, dx$$

Then ϕ_n is a bounded linear functional on $\mathbb{H}$ and can be represented by

$$\phi_n(f) = \langle e_n, f \rangle$$

where $e_n = \frac{1}{\sqrt{2\pi}} e^{inx} \in L^2(\mathbb{T})$. Therefore, from the perspective of operator norms:

$$\|\phi_n\|_{operator} = \|f_n\|_{L^2(\mathbb{T})} = 1$$

Remark 12.4. *Suppose $\phi \in \mathcal{H}^*, \phi \neq 0$, then $\phi = \phi_y$ for some $y \in (Ker(\phi))^\perp$. Then for all $z \in (Ker(\phi))^\perp$,*

$$y = \frac{\overline{\phi(z)}}{\|z\|^2} z \implies z = \frac{\|z\|^2}{\overline{\phi(z)}} y \in span\{y\} \subset (Ker(\phi))^\perp$$

Therefore,

$$(Ker(\phi))^\perp = span\{y\}$$

Returning to our example, before J is one-to-one, onto, and $\|\phi_y\| = \|y\|$. Therefore, J is an isometry between $\mathcal{H}$ and $\mathcal{H}^*$. So we may identity $\mathcal{H}$ with $\mathcal{H}^*$ and $\mathcal{H}$ is self-dual.

Remark 12.5. $\phi_{\lambda y} = \overline{\lambda} \phi_y$. *That is,*

$$\phi_{\lambda y} = \langle \lambda y, x \rangle = \overline{\lambda} \langle y, x \rangle = \overline{\lambda} \phi_y$$

$$\implies J(\lambda y) = \overline{\lambda} J(y) \implies \text{ anti-linear}$$

Conclusion: *J is isomorphic as a Banach space and anti-isomorphic as Hilbert spaces.*

12.2 Application: Existence of Adjoint Operators

Proposition 12.6. *Let $A : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator on a Hilbert space $\mathcal{H}$. There exists a unique bounded linear operator $A^* : \mathcal{H} \to \mathcal{H}$ such that*

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

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

Definition 12.7. *A^* is called the adjoint of the operator.*

Proof. (Existence) For each fixed $x \in \mathcal{H}$, then map

$$\begin{aligned} \phi_x : \mathcal{H} &\to H \\ y &\to \phi_x(y) = \langle x, Ay \rangle \end{aligned}$$

is linear and bounded. By the Riesz Representation Theorem, $\exists$ a unique $z \in \mathcal{H}$ such that

$$\phi_x(y) = \langle z, y \rangle$$

We set $A^* x = z$, then

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

∎

Proposition 12.8. *A^* is linear.*

Proof. Observe:

$$\langle A^*(ax_1+bx_2),y\rangle = \langle ax_1+bx_2,Ay\rangle = \bar{a}\langle x_1,Ay\rangle+\bar{b}\langle x_2,Ay\rangle = \bar{a}\langle A^*x_1,y\rangle+\bar{b}\langle A^*x_2,y\rangle = \langle aA^*x_1+bA^*x_2,y\rangle$$

$$\implies A^*(ax_1+bx_2) = aA^*x_1+bA^*x_2$$

∎

Proposition 12.9. *A^* is bounded.*

Proof.

$$\|A^*x\|^2 = \langle A^*x,A^*x\rangle = \langle x,AA^*x\rangle \le \|x\|\,\|AA^*x\| \le \|x\|\,\|A\|\,\|A^*x\|$$
$$\implies \|A^*x\| \le \|x\|\,\|A\|$$
$$\implies \|A^*\| \le \|A\|$$

∎

Returning to proposition 6.6, we prove uniqueness:

Proof. (Uniqueness) Assume

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

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

$$\langle A^*x-Bx,y\rangle = 0\forall y$$
$$\implies \langle A^*x-Bx,A^*-Bx\rangle = 0$$
$$\implies A^*x = Bx\forall x$$
$$\implies A^* = B$$

∎

Lemma 12.10 (Properties of Adjoint Operators)**.** *1. $A^{**}=A$*

2. $\|A^*\| = \|A\|$

3. $(AB)^* = B^*A^*$

Proof. 1. $\langle (A^*)^*x,y\rangle = \langle x,(A^*)y\rangle = \overline{\langle (A^*)y,x\rangle} = \overline{\langle y,Ax\rangle} = \langle Ax,y\rangle \implies A^{**}=A$

2. $\|A^*\| \le \|A\| = \|A^{**}\| \le \|A^*\|$

3. $\langle (AB)^*x,y\rangle = \langle x,ABy\rangle = \langle A^*x,By\rangle = \langle B^*A^*x,y\rangle$

∎

Example 12.11. *Let $\mathcal{H}=\mathbb{R}^n$. Then take a linear map on $\mathcal{H}: Ax$, where*

$$A = (a_{i,j})_{n\times m}$$

Then

$$\langle x,Ay\rangle = x^TAy = (A^Tx)^Ty = \langle A^Tx,y\rangle$$

where A^T is the transpose operation. Moreover, the adjoint of A is A^T.

Example 12.12. *Let $\mathcal{H}=\mathbb{C}^n$. Then*

$$\langle x,Ay\rangle = x^HAy = (A^Hx)^Hy = \langle A^Hx,y\rangle$$

where A^H is the hermitian transpose operation. Moreover, the adjoint of A is $A^H=\overline{A^T}$.

Example 12.13. *Let $\mathcal{H} = \ell^2(\mathbb{N})$. Recall the right-shift operator:*

$$S : (x_1, x_2, \dots) \to (0, x_1, x_2, \dots)$$

and the left-shift operator:

$$T : (x_1, x_2, x_3, \dots) \to (x_2, x_3, \dots)$$

Observe:

$$\langle x, Sy\rangle = \langle (x_1, x_2, \dots), (0, y_1, y_2, \dots)\rangle = \sum_{i=1}^{\infty} \bar{x}_{i+1} y_i = \langle (x_2, x_3, \dots), (y_1, y_2, \dots)\rangle = \langle Tx, y\rangle$$

Therefore, $S^* = T$ and $T^* = S^{**} = S$ by the previous lemma.

Example 12.14. *Let $\mathcal{H} = L^2(\mathbb{R}^n)$. Consider the operator*

$$Kf(x) = \int_{\mathbb{R}^n} k(x,y) f(y)\ dy$$

for all $f \in L^2(\mathbb{R}^n)$. Then

$$\begin{aligned}
\langle f(x), Kg(x)\rangle &= \int_{\mathbb{R}^n} \overline{f(x)} \int_{\mathbb{R}^n} k(x,y) g(y)\ dy\ dx \\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \overline{f(x)} k(x,y) g(y)\ dxdy \\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \overline{f(x)} k(x,y)\ dx\ g(y)\ dy \\
&= \int_{\mathbb{R}^n}\overline{\int_{\mathbb{R}^n} f(x) \overline{k(x,y)}\ dx}\ g(y)\ dy \\
&= \int_{\mathbb{R}^n}\overline{\int_{\mathbb{R}^n} f(y) \overline{k(y,x)}\ dy}\ g(y)\ dy && \textit{Change of variables } x \iff y \\
&= \langle K^* f, g\rangle
\end{aligned}$$

where

$$K^* f(x) = \int_{\mathbb{R}^n} f(y)\overline{k(y,x)}\ dy$$

Theorem 12.15. *If $A : \mathcal{H} \to \mathcal{H}$ is a bounded linear operator. Then*

$$\overline{\mathit{range}(A)} = (\mathit{Ker}(A^*))^{\perp}$$

$$\mathit{Ker}(A) = (\mathit{range}(A^*))^{\perp}$$

Hence, $\mathcal{H} = \overline{\mathit{range}(A)} \oplus \mathit{Ker}(A^*) = \overline{\mathit{range}(A^*)} \oplus \mathit{Ker}(A)$.

Proof. For all $x \in \text{range}(A), \exists y \in \mathcal{H}$ such that $x = Ay$. For any $z \in \text{Ker}(A^*)$

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

$$\implies \text{range}(A) \perp \text{Ker}(A^*) \implies \overline{\text{range}(A)} \subset (\text{Ker}(A^*))^{\perp}$$

On the other hand, for all $x \in (\text{range}(A))^{\perp}, \forall y \in \mathcal{H}$,

$$\langle A^* x, y\rangle = \langle x, Ay\rangle = 0 \implies A^* x = 0 \implies x \in \text{Ker}(A^*)$$

$$\implies (\text{range}(A))^{\perp} \subset \text{Ker}(A^*) \implies ((\text{range}(A))^{\perp})^{\perp} \supset (\text{Ker}(A^*))^{\perp} \implies \overline{\text{range}(A)} \supset (\text{Ker}(A^*))^{\perp}$$

We can then apply the same logic above to A^*. ∎

Application: Suppose we wish to solve the linear equation

$$Ax = y$$

where $A : \mathcal{H} \to \mathcal{H}$ is a bounded, linear operator. Then the solution exists if and only if $y \in$ range(A).

1. If range(A) is closed, the solution exists if and only if $y \in (\mathrm{Ker}(A^*))^\perp$
2. If range(A) is closed and $\mathrm{Ker}(A^*) = \{0\}$. Then the solution always has a solution for every $y \in \mathcal{H}$.

12.3 Self-Adjoint Operators

Definition 12.16. *$A \in B(\mathcal{H}, \mathcal{H})$ is self-adjoint if $A = A^*$. That is,*

$$\langle x, Ay \rangle = \langle Ax, y \rangle$$

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

Example 12.17. *Canonical Examples of Self-Adjoint Operators:*

$\mathcal{H}$	Operator	Adjoint	Self-Adjoint
$\mathbb{R}^n$	Ax	$A^T x$	$A = A^T$ (Symmetric)
$\mathbb{C}^n$	Ax	$\overline{A^T} x$	$A = A^*$ (Hermitian)
$\mathbb{L}^2(\mathbb{R}^n)$	$Kf(x) = \int_{\mathbb{R}^n} k(x,y) f(y)\ dy$	$\int_{\mathbb{R}^n} \overline{k(y,x)} f(y)\ dy$	$k(x,y) = \overline{k(y,x)}$

Example 12.18. *Let $P : \mathcal{H} \to \mathcal{H}$ be an orthogonal projection. That is, P is linear and $P^2 = P$. Then if $M = range(P)$, then*

$$\mathcal{H} = M \oplus M^\perp$$

where $M^\perp = ker(P)$. We can uniquely write $x \in \mathcal{H}$ as

$$x = px + (x - px)$$

$$\langle x, py \rangle = \langle px, py \rangle = \langle px, y \rangle \implies P^* = P$$

Example 12.19. *For any $A \in B(\mathcal{H}, \mathcal{H})$ then A^*A is self-adjoint because $(A^*A)^* = A^*A^{**} = A^*A$*

Lemma 12.20. *If A is a bounded, self-adjoint operator on a Hilbert space, then*

$$\|A\| = \sup_{\|x\|=1} |\langle x, Ax \rangle|$$

Proof. For all $x \in \mathcal{H}$ with $\|x\| = 1$, then

$$|\langle x, Ax \rangle| \le \|x\|\,\|Ax\| \le \|x\|\,\|A\|\,\|x\| = \|A\| \implies \|A\| \ge \sup_{\|x\|=1} |\langle x, Ax \rangle|$$

On the other hand, let $\alpha = \sup_{\|x\|=1} |\langle x, Ax \rangle|$. We want to show that $\|A\| \le \alpha$. That is,

$$\|Ax\| \le \alpha$$

for every $x \in \mathcal{H}$ with $\|x\| = 1$.
Case: If $Ax = 0$, then $\|Ax\| = 0 \le \alpha$.
Case: If $Ax = \alpha$, let $y = \frac{Ax}{\|Ax\|}$, then $\|y\| = 1$ and

$$\langle y, Ax \rangle = \left\langle \frac{Ax}{\|Ax\|}, Ax \right\rangle = \frac{\|Ax\|^2}{\|Ax\|} = \|Ax\|$$

So

$$\langle y, Ax\rangle = \overline{\langle y, Ax\rangle} = \langle Ax, y\rangle = \langle x, Ay\rangle$$

$$\begin{aligned}\langle x+y, A(x+y)\rangle - \langle x-y, A(x-y)\rangle &= \langle x, Ax\rangle + \langle x, Ay\rangle + \langle y, Ax\rangle + \langle y, Ay\rangle - \langle x, Ax\rangle + \langle x, Ay\rangle + \langle y, Ax\rangle - \langle y, Ay\rangle \\ &= 4\langle y, Ax\rangle = 4\|Ax\|\end{aligned}$$

Thus,

$$4\|Ax\| \le |\langle x+y, A(x+y)\rangle| + |\langle x-y, A(x-y)\rangle| \le \alpha\left(\|x+y\|^2 + \|x-y\|^2\right) \le \alpha(2\|x\|^2 + 2\|y\|^2) \le 4\alpha$$

Therefore, $\|Ax\| \le \alpha \implies \|A\| \le \alpha$. Thus, $\|A\| = \alpha$. ∎

Corollary 12.20.1. *If A is a bounded, linear operator on $\mathcal{H}$, then $\|A^*A\| = \|A\|^2$.*

Proof. Since A^*A is self-adjoint,

$$\begin{aligned}\|A^*A\| &= \sup_{\|x\|=1} |\langle x, A^*Ax\rangle| \\ &= \sup_{\|x\|=1} |\langle Ax, Ax\rangle| \\ &= \sup_{\|x\|=1} \|Ax\|^2 \\ &= \left(\sup_{\|x\|=1} \|Ax\|\right)^2 \\ &= (\|A\|)^2\end{aligned}$$

∎

Corollary 12.20.2. *If A is self-adjoint, then $\|A^2\| = \|A\|^2$.*

Corollary 12.20.3. *If P is a nonzero orthogonal projection on $\mathcal{H}$, then $\|P\| = 1$.*

Proof. $\|P^2\| = \|P\|^2 \implies \|P\| = 1$ or $\|P\| = 0$. Since P is nonzero, then $\|P\| = 1$. ∎

12.4 Unitary Operators

Definition 12.21. *A map $U : \mathcal{H}_1 \to \mathcal{H}_2$ between two (complex) Hilbert spaces is unitary if*

1. *U is linear.*
2. *U is a bijection.*
3. *U preserves the inner project*

$$\langle Ux, Uy\rangle_{\mathcal{H}_2} = \langle x, y\rangle_{\mathcal{H}_1}$$

Remark 12.22. *Condition (3) is equivalent to*

$$\|Ux\|_{\mathcal{H}_2} = \|x\|_{\mathcal{H}_1}$$

for all $x \in \mathcal{H}_1$ due to the polarization formula

$$\langle x, y\rangle = \frac{1}{4}\left[\|x+y\|^2 - \|x-y\|^2 + i\left\|\frac{x}{i} + y\right\|^2 - i\left\|\frac{x}{i} - y\right\|^2\right]$$

Definition 12.23. *Two Hilbert spaces are isomorphic if there is a unitary operator between them.*

Remark 12.24. *A linear map $U : \mathcal{H} \to \mathcal{H}$ is unitary if and only if*

$$U^*U = UU^* = I$$

Why? Observe:

$$U^*U = UU^* = I \implies U \text{ is a bijection}$$

Also,

$$\langle Ux, Uy \rangle = \langle U^*Ux, y \rangle = \langle x, y \rangle$$

Therefore, U is unitary. On the other hand,

$$\langle Ux, Uy \rangle = \langle x, y \rangle \implies \langle U^*Ux, y \rangle = \langle x, y \rangle$$

$$\implies U^*Ux = x$$

for all x. Moreover, $U^*U = I$. Now, for all $x, y \in \mathcal{H}$, since U is onto, then there exists $z \in \mathcal{H}$ sich that $x = Uz$. Therefore,

$$\langle UU^*x, y \rangle = \langle UU^*Uz, y \rangle = \langle Uz, y \rangle = \langle x, y \rangle \implies UU^* = I$$

Example 12.25. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces of the same dimension. Let $\{u_\alpha\}_{\alpha \in A}$ be an orthonormal basis of $\mathcal{H}_1$, and $\{v_\alpha\}_{\alpha \in A}$ is an orthonormal basis of $\mathcal{H}_2$. Define*

$$U : \mathcal{H}_1 \to \mathcal{H}_2$$
$$\sum_{\alpha \in A} c_\alpha u_\alpha \to \sum_{\alpha \in A} c_\alpha v_\alpha$$

Then U is unitary. Thus, Hilbert spaces of the same dimension are isomorphic.

Remark 12.26. *In particular, they are isomorphic to $\ell^2(A)$.*

Example 12.27.

$$U : L^2(\mathbb{T}) \to \ell^2(\mathbb{Z})$$
$$f = \sum_{n=-\infty}^{\infty} \hat{f}_n \frac{e^{inx}}{\sqrt{2\pi}} \to \sum_{n=-\infty}^{\infty} \hat{f}_n e_n$$

is unitary.

Example 12.28. *Define*

$$U : \mathcal{H}_1 \to \mathcal{H}_2$$
$$\sum_{\alpha \in A} c_\alpha u_\alpha \to \sum_{\alpha \in A} c_\alpha \lambda_\alpha v_\alpha$$

where $\lambda_\alpha = e^{i\theta_\alpha}$ with $\theta_\alpha \in \mathbb{R}$. To see that this mapping is unitary, we check:

- *One-to-one:*

$$\sum_{\alpha \in A} c_\alpha \lambda_\alpha v_\alpha = 0 \implies c_\alpha \lambda_\alpha \implies c_\alpha = 0$$

- *Onto:*

$$\sum_{\alpha \in A} \frac{c_\alpha}{\lambda_\alpha} \lambda_\alpha v_\alpha = 0 \implies c_\alpha \implies c_\alpha = 0$$

- *U preserves the inner product:*

$$\langle x, y\rangle = \sum_{\alpha\in A} \overline{x_\alpha} y_\alpha$$

Then

$$\langle Ux, Uy\rangle = \sum_{\alpha\in A} \overline{x_\alpha\lambda_\alpha} y_\alpha\lambda_\alpha = \sum_{\alpha\in A} \overline{x_\alpha} y_\alpha |\lambda_\alpha|^2 = \sum_{\alpha\in A} \overline{x_\alpha} y_\alpha = \langle x, y\rangle$$

Example 12.29. *(Translation on Torus $\mathbb{T}$) For all $a \in \mathbb{T}$,*

$$T_a : L^2(\mathbb{T}) \to L^2(\mathbb{T})$$
$$f \to T_a f$$

where for every $x \in \mathbb{T}, T_a f(x) = f(x-a)$. Clearly, T_a is both linear and a bijection. Lastly, to see the preservation of the inner product:

$$\langle T_a f, T_a g\rangle = \int_{\mathbb{T}} \overline{f(x-a)} g(x-a)\ dx = \int_{\mathbb{T}} \overline{f(x)} g(x)\ dx = \langle f, g\rangle$$

Example 12.30. *(The periodic Hilbert Transformation)*

$$H : L^2(\mathbb{T}) \to L^2(\mathbb{T})$$
$$f = \sum_{n=-\infty}^{\infty} \hat{f}_n \frac{e^{inx}}{\sqrt{2\pi}} \to f = \sum_{n=-\infty}^{\infty} \hat{f}_n i \cdot sign(n) \frac{e^{inx}}{\sqrt{2\pi}}$$

where

$$sign(n) = \begin{cases} -1 & n < 0 \\ 0 & n = 0 \\ 1 & n > 0 \end{cases}$$

This mapping is not one-to-one because it maps vertical shifts of functions to the same output. Specifically,

$$H(\ constant) = 0$$

So H is not unitary on $L^2(\mathbb{T})$. But we can restrict ourselves to a closed subset for which this mapping is unitary. Consider letting $\mathcal{H} = \{f \in L^2(\mathbb{T}) : \hat{f}_0 = 0\} = \{f \in L^2(\mathbb{T}) : \int_0^{2\pi} f(x)\ dx = 0\}$. Then $H : \mathcal{H} \to \mathcal{H}$ is unitary.

12.5 Weak Convergence in a Hilbert Space

Recall $x_n \rightharpoonup x$ in a normed linear space $(X, \|\cdot\|)$ if

$$\phi(x_n) \to \phi(x) \quad \forall \phi \in X^*$$

Now, suppose $\mathcal{H}$ is a Hilbert space, by Riesz representation theorem, $\mathcal{H}^* = \mathcal{H}$ in the sense that for all $\phi \in \mathcal{H}^*, \exists y \in \mathcal{H}$ such that

$$\phi(x) = \langle y, x\rangle$$

This translates weak convergence in the context of Hilbert spaces in the following way:

Definition 12.31. *A sequence $\{x_n\}_{n=1}^{\infty}$ in a Hilbert space $\mathcal{H}$ converges weakly to $x \in \mathcal{H}$ if*

$$\langle x_n, y\rangle \to \langle x, y\rangle$$

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

Proposition 12.32. *If $\{x_n\}_{n=1}^{\infty} \to x$ strongly, (i.e. $\|x_n - x\| \to 0$) in $\mathcal{H}$, then $x_n \rightharpoonup x$ in $\mathcal{H}$.*

Proof.

$$|\langle x_n, y\rangle - \langle x, y\rangle| = |\langle x_n - x, y\rangle| \leq \|x_n - x\| \, \|y\| \to 0 \cdot \|y\| = 0$$

■

Remark 12.33. *But, $x_n \rightharpoonup x \not\Longrightarrow x_n \to x$ when dim$(X) = \infty$.*

Example 12.34. *Let $\mathcal{H}$ be an infinite dimensional Hilbert space. Let $\{e_n\}_{n=1}^{\infty}$ be any orthonormal set in $\mathcal{H}$. Then*

$$e_n \rightharpoonup 0 \text{ but } e_n \not\to 0$$

Proof. For all $y \in \mathcal{H}$, by Bessel's Inequality,

$$\sum_{n=1}^{\infty} |\langle e_n, y\rangle|^2 \leq \|y\|^2 < \infty$$

$\langle e_n, y\rangle \to 0$ as $n \to \infty \implies e_n \rightharpoonup 0$.
On the other hand,

$$\|e_n - e_m\| = \sqrt{2}\delta_{m,n} \implies \{e_n\} \text{ is not Cauchy}$$

Therefore, $e_n \not\to 0$. ■

Example 12.35. *Let $\mathcal{H} = L^2([0,1])$, $\{e_n = \sqrt{2}\sin(n\pi x)\}_{n=1}^{\infty}$ is an orthonormal basis. Then $e_n \rightharpoonup 0$ but $e_n \not\to 0$. Further, for any $f \in L^2([0,1])$*

$$\langle e_n, f\rangle = \sqrt{2}\int_0^1 \sin(n\pi x)\ f(x)\ dx \to 0 \text{ as } n \to \infty$$

Notice, this is the Riemann-Lebesgue Lemma.

Question 12.36. *Given a sequence $\{x_n\}$ in a Hilbert space $\mathcal{H}$, how to prove $x_n \rightharpoonup x$?*

Brute Force Method: By definition, $\langle x_n, y\rangle \to \langle x, y\rangle$ for ALL $y \in \mathcal{H}$. This can be exhausting. Instead, we can leverage a dense set in $\mathcal{H}$.

Theorem 12.37. *Suppose that $\{x_n\}$ is a sequence in a Hilbert space $\mathcal{H}$ and D is a dense subset of $\mathcal{H}$. Then $x_n \rightharpoonup x$ if and only if:*

1. *$\|x_n\| \leq M$ for some constant M*
2. *$\langle x_n, y\rangle \to \langle x, y\rangle$ for all $y \in D$.*

Proof. ($\Leftarrow$) We want to show for all $z \in \mathcal{H}$, $\langle x_n, z\rangle \to \langle x, z\rangle$, which is equivalent to showing:

$$\langle x_n - x, z\rangle \to 0$$

We notice that if we want the following quantity to be small, it must also be true that:

$$\langle \underbrace{x_n - x}_{\text{bounded}}, \underbrace{z - y}_{\text{small}}\rangle + \langle \underbrace{x_n - x, y}_{\text{small}}\rangle$$

So for all $z \in \mathcal{H}$, and for every $\epsilon > 0$, since D is dense in $\mathcal{H}$, there must exist $y \in D$ such that

$$\|z - y\| \leq \epsilon$$

Since $\langle x_n - x, t\rangle \to 0$, then there exists N such that

$$|\langle x_n - x, y\rangle| < \epsilon$$

whenever $n \geq N$. Therefore,

$$|\langle x_n - x, z\rangle| \leq |\langle x_n - x, z - y\rangle| + |\langle x_n - x, y\rangle|$$
$$\leq (M + \|x\|)\epsilon + \epsilon$$

$$\implies \langle x_n - x\rangle \to 0 \implies x_n \rightharpoonup x$$

($\Rightarrow$) This requires something called the Uniform Boundedness principle. ■

Theorem 12.38 (Banach - Steinhaus Theorem (aka Uniform Boundedness Principle))**.** *Suppose $\{f_n\}$ is a set of linear bounded functionals on a Banach space X. If $\sup_n |f_n(x)| < \infty$, for all $x \in X$, then*

$$\sup_n \|f_n\| < \infty$$

Alternative Interpretations:

- If $\{f_n(x)\}$ is a point-wise bounded sequence, then $\{\|f_n\|\}$ is uniformly bounded.
- If for all x, there exists M_x such that
$$|f_n(x)| \leq M_x \|x\|$$
for all n, then there exists $M > 0$ such that
$$|f_n(x)| \leq M \|x\|$$
for all n, for all x.

Corollary 12.38.1. *If $\sup_n \|f_n\| = \infty$, then there exists $x_0 \in X$ such that*

$$\sup_n |f_n(x_0)| = \infty$$

Proposition 12.39. *$y_n \rightharpoonup y$ in $\mathcal{H} \implies \{y_n\}$ is bounded.*

Proof. Since $y_n \in H = H^*$ we let $f_n = f_{y_n}$. That is,

$$f_n(x) = \langle y_n, x\rangle \forall x \in \mathcal{H}$$

Notice, $\|f_n\| = \|y_n\|$ for any $x \in \mathcal{H}$. Since $\{f_n(x)\}$ is convergent to $\langle y, x\rangle$ if is a bounded sequence. By the uniform boundedness principle, $\{\|f_n\|\}$ is bounded and therefore, $\{\|y_n\|\}$ is bounded. ■

Proposition 12.40 (Lower Semicontinuity of Weak Convergence)**.** *If $x_n \rightharpoonup x$, then $\|x\| \leq \liminf_{n\to\infty} \|x_n\|$*

Proof.
$$\|x\|^2 = \langle x, x\rangle = \lim_{n\to\infty} \langle x_n, x\rangle \leq \liminf_{n\to\infty} \|x_n\| \|x\|$$

■

In the theorem, D is any arbitrary dense subset of $\mathcal{H}$. In particular, setting $D = \text{span}\{e_\alpha\}$ where $\{e_\alpha\}$ is an orthonormal basis of $\mathcal{H}$. (Not necessarily countable)
In this case, the condition $\langle x_n, y\rangle \to \langle x, y\rangle$ for all $y \in D$ becomes: $\langle x_n, y\rangle \to \langle x, y\rangle$ for all $y \in \text{span}\{e_\alpha\}$ which is equivalent to

$$\langle x_n, e_\alpha\rangle \to \langle x, e_\alpha\rangle$$

Therefore, $x_n \rightharpoonup x$ in $\mathcal{H}$ if and only if

1. $\{x_n\}$ is bounded AND
2. $\langle x_n, e_\alpha\rangle \to \langle x, e_\alpha\rangle$ for all α

Remark 12.41. *Condition 2 alone does NOT imply $x_n \rightharpoonup x$.*

Example 12.42. *Let $\mathcal{H} = \ell^2(\mathbb{N})$ and e_n the canonical vector. Let*

$$x_n = ne_n$$

Then for each coordinate e_m

$$\langle x_n, e_m \rangle = m\delta_{n,m} \to 0$$

as $n \to \infty$. On the other hand, for all $y \in \ell^2(\mathbb{N})$,

$$\langle x_n, y \rangle = ny_n$$

which does not necessarily converge to zero! Consider

$$y = (1, 1/2, 1/3, \ldots, 1/n, \ldots)$$

Then $y \in \ell^2(\mathbb{N})$ because $\|y\|_2^2 = \sum_n \frac{1}{n^2} < \infty$ by

$$\langle x_n, y \rangle = n \cdot \frac{1}{n} = 1 \not\to 0$$

Therefore, $x_n \not\rightharpoonup 0$ even though $\{x_n\}$ is bounded.

Now we consider some typical ways that a weakly convergent sequence fails to converge strongly by constructing $f_n \rightharpoonup 0$ by $\|f_n\| = 1$.

1. Oscillation

   Example 12.43. *Define $H = L^2([0,1])$, $e_n = \sqrt{2\pi}\sin(n\pi x)$ Then for all n, $\{e_n\}_{n=1}^{\infty}$ is an orthonormal basis of $\mathcal{H}$ with $e_n \rightharpoonup 0$ but $e_n \not\to 0$. Now, for all $f \in L^2([0,1])$*

$$\langle e_n, f \rangle = \sqrt{2\pi} \int_0^1 f(x)\sin(n\pi x)\ dx \to 0$$

2. Concentration of Mass

   Example 12.44. *Define $H = L^2([0,1])$, and define*

$$f_n = \begin{cases} \sqrt{n} & x < \frac{1}{n} \\ 0 & x > \frac{1}{n} \end{cases}$$

   Clearly, $\|f_n\|_2 = 1$. Now, for any $f \in C^\infty([0,1])$, we see:

$$\begin{aligned} \langle f_n, f \rangle &= \int_0^{\frac{1}{n}} \sqrt{n} f(x)\ dx \\ &= \sqrt{n} \int_0^{\frac{1}{n}} f(x)\ dx && = \sqrt{n} f(x_n) \frac{1}{n} \text{Mean-Value Theorem} \\ &= \frac{f(x_n)}{\sqrt{n}} \to \frac{f(0)}{\infty} = 0 \end{aligned}$$

   Since $\|f_n\|_2 = 1$ is bounded and $C^\infty[0,1]$ is dense in $\mathcal{H}$, then we conclude $f_n \rightharpoonup 0$

3. Mass escaping to infinity Let $H = L^2(\mathbb{R})$

$$f_n = \mathcal{X}_{[n,n+1]}$$

Notice, $\|f_n\| = 1$ for all $f \in C_c^\infty(\mathbb{R}) \subset L^2(\mathbb{R})$ a dense subset. Observe:

$$\begin{aligned}\langle f_n, f\rangle &= \int_n^{n+1} f(x)\ dx \\ &= \int_{\text{supp}(f)\cap[n,n+1]} f\ dx \to 0\end{aligned}$$

Therefore, $f_n \rightharpoonup 0$.

Proposition 12.45. *$x_n \to x$ if and only if $x_n \rightharpoonup x$ and $\|x\| = \lim_{n\to\infty} \|x_n\|$*

Proof. ($\Rightarrow$) $x_n \to x \implies x_n \rightharpoonup x$. Also

$$|\,\|x\| - \|x_n\|\,| \leq \|x - x_n\| \to 0 \implies \lim_{n\to\infty} \|x_n\| = \|x\|$$

($\Leftarrow$) If $x_n \rightharpoonup x$ and $\|x_n\| \to \|x\|$, then

$$\|x_n - x\|^2 = \|x_n\|^2 + \|x\| - \langle x_n, x\rangle - \langle x, x_n\rangle \to \|x\|^2 + \|x\|^2 - \langle x, x\rangle - \langle x, x\rangle = 0$$

∎

12.6 Compactness

Proposition 12.46. *Every bounded sequence in a Hilbert space has a weakly convergent subsequence. That is, if $\|x_n\| \leq M$, for all n, then there exists a subsequence $\{x_{n_k}\}_k$ such that $x_{n_k} \rightharpoonup x$ for some $x \in \mathcal{H}$.*

Remark 12.47. *Let $\{x_n\}$ be a bounded sequence in $\mathcal{H}$.*

1. *If dim$(\mathcal{H}) < \infty$, then $\mathcal{H} = \mathbb{C}^n$. By Heine-Borel theorem, $\{x_n\}$ has a (strongly) convergent subsequence.*
2. *If dim$(\mathcal{H}) = \infty$, then $\{x_n\}$ may fail to have a strongly convergent subsequence. i.e. an orthonormal sequence.*
3. *This proposition says $\{x_n\}$ always has a weakly convergent subsequence even when dim$(\mathcal{H}) = \infty$.*

Corollary 12.47.1. *Let A be a subset of a Hilbert space $\mathcal{H}$. Then A is bounded if and only if A is weakly precompact.*

Proof. (of Corollary)
($\Rightarrow$) Suppose A is bounded. Than any sequence $\{x_n\}$ in A is bounded, and thus has a weakly convergent subsequence!
($\Leftarrow$) Assume that A is weakly precompact. That is, every sequence in A has a weakly convergent subsequence. We also assume A is unbounded. Then there exists a sequence $\{x_n\}$ in A such that

$$\|x_n\| \geq \max_{1\leq i\leq n-1} \{\|x_i\|, n\}$$

Thus $\{x_n\}$ has no bounded subsequence. Since A is precompact, $\{x_n\}$ must have a weakly convergent subsequence $\{x_{n_k}\}$, so $\{x_{n_k}\}$ is bounded in $\mathcal{H} \implies$ A contradiction! ∎

Proof. (of Proposition)
We'll prove the result only for a separable Hilbert space. Without loss of generality, we may assume that $\mathcal{H} = \ell^2(\mathbb{N})$. Then

$$x_n = (x_n^{(1)}, x_n^{(2)}, \ldots) \in \ell^2(\mathbb{N})$$

with

$$\|x_n\|^2 = \sum_{k=1}^{\infty} |x_n^{(k)}|^2 \leq M \implies |x_n^{(k)}| \leq M, \forall n, \forall k$$

We now use the diagonal argument to find a weakly convergent subsequence of $\{x_n\}$ since $(x_n^{(1)})_{n=1}^{\infty}$ is a bounded sequence in $\mathbb{C}$, so it has a convergent subsequence $(x_{1,n}^{(1)})_{n=1}^{\infty} \to x^{(1)} \in \mathbb{C}$ Further, since $(x_n^{(1)})_{n=1}^{\infty}$ is bounded sequence in $\mathbb{C}$, it has a convergent subsequence $(x_{2,n}^{(2)})_{n=1}^{\infty} \to x^{(2)}$. So we see, continuing this process, we get the subsequences:

$$(x_n^{(1)})_{n=1}^{\infty} \to (x_{1,n}^{(1)})_{n=1}^{\infty} \to x^{(1)}$$
$$(x_n^{(1)})_{n=1}^{\infty} \to (x_{2,n}^{(2)})_{n=1}^{\infty} \to x^{(2)}$$
$$\vdots$$
$$(x_n^{(1)})_{n=1}^{\infty} \to (x_{k,n}^{(k)})_{n=1}^{\infty} \to x^{(k)}$$
$$\vdots$$

We now take the diagonal elements to form a subsequence

$$(x_{n,n}^{(k)})_{n=1}^{\infty} \to x^{(k)}$$

as $n \to \infty$ for each $k = 1, 2, \ldots$. Construct $x^* = (x^{(1)}, x^{(2)}, \ldots)$. We want to see that this sequence $x^* \in \ell^2$. Observe:

$$\begin{aligned}
\|x^*\|^2 &= \sum_{k=1}^{\infty} |x^{(k)}|^2 \\
&= \sum_{k=1}^{\infty} \lim_{n\to\infty} |x_{n,n}^{(k)}|^2 \\
&\leq \liminf_{n\to\infty} \sum_{k=1}^{\infty} |x_{n,n}^{(k)}|^2 && \text{Fatou's Lemma} \\
&= \lim_{n\to\infty} \|x_{n,n}\|^2 \leq M^2
\end{aligned}$$

Therefore, $x^* \in \ell^2(\mathbb{N})$ and $x_{n,n}^k \to x^{(k)}$ for each k. Moreover, $x_{n,n} \rightharpoonup x^*$. ■

Corollary 12.47.2 (Banach-Alaoglu)**.** *The closed unit ball of a Hilbert space is (sequentially) weakly compact.*

Proof. For any sequence $\{x_n\}$ in $\overline{B_1(0)} \subset \mathcal{H}$, then $\|x_n\| \leq 1$. Thus, $\{x_n\}$ has a subsequence $\{x_{n_k}\} \rightharpoonup x$ for some $x \in \mathcal{H}$. We need to check that $x \in \mathcal{H}$ to demonstrate this is weakly precompact. Since

$$\|x\| \leq \liminf_{R\to\infty} \|x_{n_k}\| \leq 1$$

we have $x \in \overline{B_1(0)}$. Therefore, $\overline{B_1(0)}$ is sequentially weakly compact, which also implies $\overline{B_1(0)}$ is weakly compact. ■

Remark 12.48. *In a metric space, compact $\equiv$ precompact and closed $\equiv$ totally bounded and complete. In a Hilbert space, weakly precompact and weakly closed = weakly compact $\equiv$ bounded and weakly closed Therefore,*

$$\text{totally bounded} \implies \text{bounded}$$

$$\text{strongly closed} \implies \text{weakly closed}$$

Proof. Suppose K is weakly closed. Then for any (x_n) in K, with $x_n \to x$ we also have $x_n \rightharpoonup x$. Since K is weakly closed, we have $x \in K$. Therefore, K is closed. ■

<u>Application</u> Consider a minimization problem

$$\min_{x \in K} f(x)$$

where $f : K \to \mathbb{R}$.

1. By the previous remark, if K is a compact metric space and f lower-semicontinuous, then the minimization problem has a minimizer.

2.

   Theorem 12.49. *If K is weakly compact and f is weakly lower semi-contiuous in the sense that*

   $$f(x) \leq \liminf_{n \to \infty} f(x_n)$$

   for every sequence (x_n) in K such that $x_n \rightharpoonup x$. Then the minimization problem has a minimizer.

   Proof. (Direct method of Calculus of Variations)

   (a) Take a minimizing sequence (x_n) in K such that

   $$f(x_n) \to \inf_{x \in K} f(x)$$

   (b) Applying the compactness result, since K is weakly bounded, (x_n) has a weakly convergent subsequence (x_{n_k}) such that $x_{n_k} \rightharpoonup x^*$ for some $x^* \in K$ since K is weakly closed.

   (c) To obtain lower semi-continuity, since f is weakly lower semi-continuous

   $$f(x^*) \leq \liminf_{k \to \infty} f(x_{n_k}) = \inf_{x \in K} f(x)$$

   Therefore, $f(x^*) = \inf_{x \in K} f(x)$ and

   $$f(x^*) = \min_{x \in K} f(x)$$

   ■

Now suppose $K \subset \mathcal{H}, f : K \to \mathbb{R}$

1. K is compact $\implies$ K is weakly compact.

2. f is lower semi-continuous $\Leftarrow$ f is weakly lower-semi continuous

Question 12.50. *Given a bounded and closed set, how can we get weak compactness?*

Theorem 12.51. *Suppose K is (strongly) closed, bounded, <u>convex</u> subset of a Hilbert space and $f : K \to \mathbb{R}$ is (strongly) lower-semicontinuous, <u>convex</u> function on K. Then the minimization problem has a minimizer. If also, f is strictly convex, then the minimizer is unique.*

Definition 12.52. *A convex set is a subset C such that for any $x, y \in C$, $0 \leq t \leq 1$*

$$tx + (1-t)y \in C$$

On the other hand, a function $f : C \to \mathbb{R}$ is convex on a convex set C if

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

Further, f is strictly convex if

$$f(tx + (1-t)y) < tf(x) + (1-t)f(y)$$

for all $x \neq y \in C, 0 < t < 1$.

Definition 12.53. *Given $\{x_1, x_2, \ldots, x_n\}$ of a linear space X, we define a the set*

$$C = \left\{ y = \sum_{k=1}^{n} \lambda_k x_k : \lambda_k \geq 0, \sum_{k=1}^{n} \lambda_k = 1 \right\}$$

each y is called a convex combination of $\{x_1, x_2, \ldots, x_n\}$.

By Mazur's lemma, if $x_n \rightharpoonup x$, then there exists a sequence (y_n) of finite convex combinations of (x_n) such that $y_n \to x$ strongly.

Proof. (Of Theorem 6.51)
Idea:

1. We show closed and convex implies weakly closed and bounded which implies weakly compact
2. f is lower semi-continuous and convex which implies f is lower semi-continuous.

1. If $x_n \rightharpoonup x$ in K, by Mazur's lemma, there exists a sequence (y_n) of finite convex combinations of (x_n) such that $y_n \to x$. Since K is convex, $y_n \in K$, Since K is closed and $y_n \to x$, it must follow that $x \in K$. Therefore, K is weakly closed. Since K is also bounded, we have proven K is weakly compact.
2. This was proven on the assigned homework.
3. Now, suppose f is strictly convex. We want to show the uniqueness of the minimizer. Assume f achieves its minimum at both x^* and y^* in K with $x^* = y^*$. Then

$$f\left(\frac{x^* + y^*}{2}\right) < \frac{1}{2}f(x^*) + \frac{1}{2}f(y^*) = f(x^*)$$

and $\frac{x^*+y^*}{2} \in K$ Contradiction! Therefore, the minimizer is unique!

■

Question 12.54. *What is K is unbounded? e.g. $K = \mathcal{H}$ or the upper-half space.*

Then we also need f to be coercive with

$$\lim_{\|x\| \to \infty} f(x) = \infty$$

Theorem 12.55. *Suppose*

1. *K is closed, convex subset of $\mathcal{H}$*
2. *$f : K \to \mathbb{R}$ is coercive, lower semi-continuous convex function.*

Then f achieves its minimum on K.

Proof. Let $x_0 \in K$, since f is coercive, there exists R such that

$$f(x) > f(x_0) + 1$$

when $\|x\| \geq R$. Now $K \cap \{x : \|x\| \leq R\}$ is closed, bounded a convex subset of $\mathcal{H}$. Therefore, a minimizer of f on this subset is a global minimizer of f on K. ■

13 The Spectrum of Bounded Linear Operators

In linear algebra, we studied eigenvalues of matrices. Here, in Banach / Hilbert spaces, we study the spectrum of bounded linear operators.

Definition 13.1. *Let A be a bounded linear operators on a Banach space X. The resolvent set of A is given by*

$$\begin{aligned}\rho(A) &:= \{\lambda \in \mathbb{C} : \lambda I - A : X \to X \text{ is invertible}\} \\ &= \{\lambda \in \mathbb{C} : (\lambda I - A)^{-1} \in B(X)\} \\ &= \{\lambda \in \mathbb{C} : \lambda I - A : X \to X \text{ is one-to-one and onto}\}\end{aligned}$$

Note 13.2. *Recall, the Banach bounded inverse theorem, which states that given a $T : X \to X$ a bounded bijection implies T^{-1} is also bounded!*

Definition 13.3. *The spectrum of A is the set*

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

Example 13.4. *Let A be a finite dimensional Banach space, $A \in B(X)$, and $\lambda I - A$ is not invertible.*
$\iff$ *$\exists$ a nonzero element $x_0 \in X$ such that $(\lambda I - A)x_0 = 0$*

$$\iff Ax_0 = \lambda x_0, x_0 \neq 0$$

λ is an eigenvalue of A. If this case,

$$\sigma(A) := \{\text{eigenvalues of } A\}$$

In general, for any $A \in B(X), \sigma(A) = \{\lambda \in \mathbb{C} : \lambda I - A \text{ is not a bigjection}\}$ which can be expressed as

$$\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$$

1. $\sigma_p(A)$ is called the point spectrum of A.

$$\begin{aligned}\sigma_p(A) :=& \{\lambda \in \mathbb{C} : \lambda I - A \text{ is not one-to-one}\} \\ &= \{\lambda \in \mathbb{C} : \mathrm{Ker}(\lambda I - A) \neq \{0\}\} \\ &= \{\lambda \in \mathbb{C} : \exists x_0 \neq 0 \text{ such that } (\lambda I - A)x_0 = 0\} \\ &= \{\lambda \in \mathbb{C} : Ax_0 = \lambda x_0 \text{ for some } x_0 \neq 0\}\end{aligned}$$

   Each $\lambda \in \sigma_p(A)$ is called an eigenvalue of A.

   Example 13.5. *When $dim(X) < \infty, \sigma(A) = \sigma_p(A)$*

2. $\sigma_c(A)$ is called the continuous spectrum of A.

$$\sigma_p(A) := \{\lambda \in \mathbb{C} : \lambda I - A \text{ is one-to-one but not onto but } \overline{\mathrm{range}(\lambda I - A)} = X\}$$

   That is, the range$(\lambda I - A)$ is dense in X.

3. $\sigma_r(A)$ is called the residual spectrum of A

$$\sigma_r(A) := \{\lambda \in \mathbb{C} : \lambda I - A \text{ is one-to-one but not onto and } \overline{\mathrm{range}(\lambda I - A)} \neq X\}$$

Example 13.6. *Let* $X = (C[0,1], \|\cdot\|_\infty)$,

$$\begin{aligned} A &: X \to X \\ f(x) &\to xf(x) \end{aligned}$$

Claim: $A \in B(X)$ *and* $\sigma(A) = \sigma_r(A) = [0,1]$.

Note 13.7. *A is linear and*

$$\|Af\|_\infty = \|xf(x)\|_\infty = \sup_{x\in[0,1]} |xf(x)| \le \|f\|_\infty \implies \|A\| \le 1$$

Moreover, since $\|A1\|_\infty = \|x\|_\infty = 1 = \|1\|_\infty \implies \|A\| = 1$ *and* $A \in B(X)$

Question 13.8. *Now,how do we find the spectrum of this operator?*

Notice, for any $\lambda \in \mathbb{C}$, *and any* $f \in C[0,1]$

$$(\lambda I - A)f = \lambda f(x) - xf(x) = (\lambda - x)f(x)$$

1. *If* $(\lambda I - A)f = 0$, *then* $(\lambda - x)f(x) = 0$ *for all* $x \in [0,1] \implies f(x) = 0$ *when* $x \neq \lambda$. *Since f is continuous, it follows that* $f = 0$. *Therefore,* $\lambda I - A$ *is one-to-one for any* $\lambda \in \mathbb{C}$. *Thus,* $\sigma_p(A) = \emptyset$.

2. *For every* $\lambda \notin [0,1]$, *we claim that* $\lambda I - A$ *is onto. Indeed, for every* $g \in C[0,1]$, *let* $f(x) = \frac{g(x)}{\lambda - x} \in C[0,1]$ *then*
$$(\lambda I - A)f(x) = (\lambda - x)f(x) = g(x)$$
Therefore, $\lambda I - A$ *is onto, therefore* $\lambda \in \rho(A)$. *Thus,* $\sigma(A) \subset [0,1]$

3. *Lastly, we need to prove* $[0,1] \subset \sigma_r(A)$. *That is, for every* $\lambda \in [0,1]$, *we want to show* $\overline{\text{range}(\lambda I - A)} \neq C([0,1])$. *Indeed, for any* $\lambda \in [0,1]$, *and any* $g \in \text{range}(\lambda I - A)$, *then there exists an* $f \in C([0,1])$ *such that*
$$g(x) = (\lambda I - A)f(x) = (\lambda - x)f(x)$$
In particular, $g(\lambda) = 0$. *Also, if* $h \in \overline{\text{range}(\lambda I - A)}$ *then there exists* $g_n \in \text{range}(\lambda I - A)$ *such that* $g_n \to h$ *uniformly on* $[0,1]$.
$$h(\lambda) = \lim_n g_n(\lambda) = 0$$
Therefore, $1 \notin \overline{\text{range}(\lambda I - A)}$. *Thus,* $\lambda \in \sigma_r(A)$. *Thus,*
$$[0,1] \subset \sigma_r(A) \subset \sigma(A) \subset [0,1] \implies \sigma(A) = \sigma_r(A) = [0,1]$$

Example 13.9. *Let* $\mathcal{H} : L^2([0,1])$

$$\begin{aligned} A &: X \to X \\ f(x) &\to xf(x) \end{aligned}$$

Notice, $L^2([0,1])$ *and* $C([0,1])$ *have different topologies, and hence the closure of range*$(\lambda I - A)$ *is different.*

Claim: $A \in B(\mathcal{H})$ *and* $\sigma(A) = \sigma_c(A) = [0,1]$.

1. *A is clearly linear and for all* $f \in \mathcal{H}$,
$$\|Af\|_2 = \|xf(x)\|_2 = \sqrt{\int_0^1 x^2 f(x)^2 \, dx} \le \sqrt{\int_0^1 f^2 \, dx} = \|f\|_2$$
Therefore, $\|A\| \le 1$, *and* $A \in B(\mathcal{H})$.

2. *For all $\lambda \in \mathbb{C}$, if $f \in Ker(\lambda I - A)$, then*

$$(\lambda I - A)f = (\lambda - x)f(x) = 0 \implies f(x) = 0 \text{ when } x \neq \lambda$$

So $f = 0$ in $L^2([0,1])$ and thus, $Ker(\lambda I - A) = \{0\} \implies A$ is one-to-one. So $\sigma_p(A) = \emptyset$.

3. *Want to show $\sigma_c(A) \subset [0,1]$. We claim that $\lambda I - A$ is onto. Indeed, for all $g \in L^2([0,1])$, and let $f(x) = \frac{g(x)}{\lambda - x} \in L^2([0,1])$ and $g = (\lambda I - A)f$.*

4. *We want to show that $[0,1] \subset \sigma_c(A)$. Notice, for all $\lambda \in [0,1]$, $\lambda I - A$ is not onto. Otherwise, $1 = (\lambda I - A)f(x) = (\lambda - x)f(x)$ for some $f(x) \in L^2([0,1])$. Then $f(x) = \frac{1}{\lambda - x} \notin L^2([0,1])$.*
Now we want to show that $\overline{range(\lambda I - A)} = L^2([0,1])$. Observe, for all $f \in L^2([0,1])$, define:

$$f_n(x) = \begin{cases} f(x) & |x - \lambda| \geq \frac{1}{n} \\ 0 & |x - \lambda| < \frac{1}{n} \end{cases}$$

Then $f_n \to f$ in L^2 because

$$\|f_n - f\|_2^2 = \int_{|x-\lambda|<\frac{1}{n}} |f(x)|^2 \, dx \to 0 \text{ as } n \to \infty$$

Also, $f_n(x) = (\lambda - x)g_n(x) \in range(\lambda I - A)$ for

$$g_n(x) = \begin{cases} \frac{f(x)}{\lambda - x} & |x - \lambda| \geq \frac{1}{n} \\ 0 & |x - \lambda| < \frac{1}{n} \end{cases}$$

with

$$\|g\|_2^2 = \int_{|x-\lambda|<\frac{1}{n}} \frac{|f(x)|^2}{|\lambda - x|^2} \, dx \leq n^2 \int_0^1 |f(x)|^2 \, dx = n^2 \|f\|_2^2 < \infty$$

So $\|g\|_2 \leq n \|f\|_2$ and $g \in L^2([0,1])$. Therefore, $f \in \overline{range(\lambda I - A)}$ and $\lambda \in \sigma_c(A)$.

Therefore, by the claims above, we arrive at the chain:

$$[0,1] \subset \sigma_c(A) \subset \sigma(A) \subset [0,1]$$

Therefore, $\sigma(A) = \sigma_c(A) = [0,1]$!

Question 13.10. *In these examples, $\sigma(A) = [0,1]$, a compact subset of $\mathbb{C}$. How about other A?*

Ans: In general, $A \in B(X) \implies \sigma(A)$ is a nonempty compact subset of $\mathbb{C}$.

Definition 13.11. *Let X be a Banach space and $A \in B(X)$. Then*

$$\rho(A) = \{\lambda \in \mathbb{C} : \lambda I - A \text{ is invertible}\}$$

For all $\lambda \in \rho(A)$, let

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

be the <u>resolvent</u> of A at λ.

Proposition 13.12. *$\rho(A)$ is an open subset of $\mathbb{C}$. Moreover, $\sigma(A)$ is closed.*

Proof. For all $\lambda_0 \in \rho(A)$, that is $R_{\lambda_0} = (\lambda_0 I - A)^{-1}$ exists, then

$$\begin{aligned}\lambda I - A &= (\lambda_0) + (\lambda - \lambda_0)I \\ &= (\lambda_0 - A)(I - (\lambda_0 - \lambda)(\lambda_0 I - A)^{-1}) \\ &= (\lambda_0 - A)(I - K)\end{aligned}$$

where $K = (\lambda_0 - \lambda)R_{\lambda_0} \in B(X)$. When $\|K\| < 1$, then

$$|\lambda_0 - \lambda| < \frac{1}{\|R_{\lambda_0}\|}$$

so $I - K$ has an inverse and

$$(I-K)^{-1} = \sum_{n=0}^{\infty} K^n = \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n R_{\lambda_0}^n$$

So $(\lambda I - A)^{-1} = (I-K)^{-1}(\lambda_0 I - A)^{-1} = \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n R_{\lambda_0}^{n+1}$. So $\lambda \in \rho(A)$ whenever $|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}\|}$ Therefore, $\rho(A)$ is open. Moreover,

$$R_\lambda = \sum_{n=0}^{\infty} R_{\lambda_0}^{n+1} (\lambda_0 - \lambda)^n$$

when $|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}\|}$ ∎

<u>Recall:</u> $f : \Omega \subset \mathbb{C} \to \mathbb{C}$ is analytic / holomorphic if $f'(z)$ exists for all $z \in \Omega$. Equivalently, if

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

in a neighborhood of z_0 for all $z_0 \in \Omega$. So, as an operator-valued function of $\lambda \in \rho(A)$, R_λ is analytic in $\rho(A)$.

Proposition 13.13. *For all $A \in B(X)$, $\sigma(A) \neq \emptyset$.*

Proof. If not, $\rho(A) = \mathbb{C}$. For all λ with $|\lambda| \geq \|A\|$

$$R_\lambda - (\lambda I - A)^{-1} = \lambda^{-1}(I - \frac{1}{\lambda}A)^{-1} = \frac{1}{\lambda}\sum_{n=0}^{\infty}\left(\frac{A}{\lambda}\right)^n = \sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} A^n$$

which is a Laurent Series Expansion. Thus,

$$\|R_\lambda\| \leq \sum_{n=0}^{\infty} \frac{\|A\|^n}{|\lambda|^{n+1}} = \frac{1}{|\lambda|}\frac{1}{1 - \frac{\|A\|}{|\lambda|}} = \frac{1}{|\lambda| - \|A\|} \to 0 \text{ as } |\lambda| \to \infty$$

Consequently, $\|R_\lambda\|$ is bounded on $\mathbb{C}$. Now, for any $f \in B(X)^*$, we consider the complex-valued function

$$\begin{aligned}g_f(\lambda) &:= f(R_\lambda) \\ &= f\left(\sum_{n=0}^{\infty}(\lambda_0 - \lambda)^n (R_{\lambda_0})^{n+1}\right) \\ &= \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n f((R_{\lambda_0})^{n+1}) && \text{since } f \text{ is linear}\end{aligned}$$

Therefore, $g_f(\lambda)$ is analytic on $\mathbb{C}$ and

$$|g_f(\lambda)| = |f(R_\lambda)| \leq \|f\|\,\|R_\lambda\| \leq \|f\|\,M < \infty$$

So $g_f(\lambda)$ is bounded on $\mathbb{C}$. By Liouville's theorem, we know that $g_f(\lambda)$ is a constant function c_f. Therefore,

$$f(R_\lambda) = c_f$$

for every $f \in B(X)^*$. Recall, for every $x_0, x_1 \in X$ with $x_0 \neq x_1$ then we can always identity an $f \in X^*$ such that $f(x_0) \neq f(x_1)$. Here, if $R_{\lambda_1} \neq R_{\lambda_2} \in B(X)$, then there must exists $f \in B(X)^*$ such that $f(R_{\lambda_1} \neq f(R_{\lambda_2})$. Therefore, R_λ is a constant operator. Since $\|R_\lambda\| \to 0$ as $|\lambda| \to \infty$, it must follow that $R_\lambda = 0$. This is impossible since $R_\lambda = (\lambda I - A)^{-1}$. ■

Proposition 13.14. *For every $A \in B(X)$, we have $\rho(A) \supset \{\lambda \in \mathbb{C} : |\lambda| > \|A\|\}$ and $\sigma(A) \subset \{\lambda \in \mathbb{C} : |\lambda| \leq \|A\|\}$. ($\sigma(A)$ is a nonempty, bounded and closed subset of $\mathbb{C}$.)*

Proof. For all $\lambda \in \mathbb{C}$ with $|\lambda| > \|A\|$. Then

$$\lambda I - A = \lambda(I - \frac{A}{\lambda})$$

is invertible because

$$\left\| \frac{A}{\lambda} \right\| = \frac{\|A\|}{|\lambda|} < 1$$

Therefore, $\lambda \in \rho(A)$. ■

Definition 13.15. *For any $A \in B(X)$, the number $r(A) = \sup_{\lambda \in \sigma(A)} |\lambda|$ is called the spectral radius of A.*

Note 13.16. $r(A) \leq \|A\|$.

Proposition 13.17. *For any $A \in B(X)$,*

$$r(A) = \lim_{n \to \infty} \|A^n\|^{\frac{1}{n}}$$

Proof. First we note that by subnormativity, $\|AB\| \leq \|A\| \|B\|$. Therefore,

$$\|A^n\| \leq \|A\|^n \implies \|A^n\|^{\frac{1}{n}} \leq \|A\|$$

- *Claim:* $r(A) \leq \liminf_{n \to \infty} \|A^n\|^{\frac{1}{n}}$
  For all $\lambda \in \sigma(A)$, we have $\lambda^n \in \sigma(A^n)$. Thus, $|\lambda^n| \leq \|A^n\| \implies |\lambda| \leq \|A^n\|^{\frac{1}{n}} \implies r(A) \leq \|A^n\|^{\frac{1}{n}} \implies r(A) \leq \liminf_{n \to \infty} \|A^n\|^{\frac{1}{n}}$.

- *Claim:* $r(A) \geq \limsup_{n \to \infty} \|A^n\|^{\frac{1}{n}}$
  Consider

  $$\begin{aligned} F : \rho(A) &\to B(X) \\ \lambda &\to R_\lambda = (\lambda I - A)^{-1} \end{aligned}$$

  Notice, $F(\lambda)$ is analytic on $\rho(A) \supset \{\lambda | |\lambda| > r(A)\}$. So F has a Laurent series expansion

  $$F(\lambda) = \sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} A^n$$

  When $|\lambda| > r(A)$, then

  $$\frac{1}{|\lambda|^{n+1}} \|A^n\| \leq M$$

  when n large enough. So

  $$\|A^n\|^{\frac{1}{n}} \leq M^{\frac{1}{n}} |\lambda|^{\frac{n+1}{n}} \to |\lambda|$$

  Therefore, $\limsup_{n \to \infty} \|A^n\|^{\frac{1}{n}} \leq |\lambda| \leq r(A)$.

Therefore, we conclude $r(A) = \lim_{n\to\infty} \|A^n\|^{\frac{1}{n}}$. ■

Now, let $X = \mathcal{H}$ a Hilbert space.

Proposition 13.18. *Suppose $A \in B(\mathcal{H})$ for some Hilbert space $\mathcal{H}$. Then for all $\lambda \in \sigma_r(A)$, we have $\overline{\lambda} \in \sigma_p(A^*)$*

Proof. For all $\lambda \in \sigma_r(A)$

$$\begin{aligned}
&\implies \overline{\operatorname{ran}(\lambda I - A)} \neq \mathcal{H} \\
&\implies (\operatorname{Ker}(\lambda I - A)^*)^{\perp} \neq \mathcal{H} \\
&\implies \operatorname{Ker}(\lambda I - A)^* \neq \{0\} \\
&\implies \operatorname{Ker}(\overline{\lambda} I - A)^* \neq \{0\} \\
&\implies \overline{\lambda} \in \sigma_p(A^*)
\end{aligned}$$

■

Proposition 13.19. *Let A be a self-adjoint bounded linear operator on a Hilbert space $\mathcal{H}$. Then*

1. $\sigma_p(A) \subset \mathcal{R}$
2. $\sigma_r(A) = \emptyset$
3. $\sigma_c(A) \subset \mathcal{R}$

Therefore, $\sigma(A) = \sigma_p(A) \cup \sigma_c(A)$ is contained in the interval $[-\|A\|, \|A\|] \subset \mathbb{R} \subset \mathbb{C}$.

Proof. 1. For all $\lambda \in \sigma_p(A)$, then there exists a nonzero x such that

$$Ax = \lambda x$$

Since A is self-adjoint,

$$\langle x, Ax\rangle = \langle Ax, x\rangle \implies \langle x, \lambda x\rangle = \langle \lambda x, x\rangle \implies \lambda \|x\|^2 = \overline{\lambda} \|x\|^2 \implies \lambda = \overline{\lambda}$$

2. Assume $\sigma_r(A) \neq \emptyset$. Then there exists $\lambda \in \sigma_r(A)$. So

$$\overline{\lambda} \in \sigma_p(A^*) = \sigma_p(A)$$

Therefore, $\overline{\lambda} \in \mathbb{R}$ and $\lambda = \bar{\lambda} \in \sigma_r(A) \cap \sigma_p(A) = \emptyset$. Contradiction! Therefore, $\sigma_r(A) = \emptyset$.

3. Assume $\exists \lambda = a + b_i \in \sigma_c(A)$ with $b \neq 0$. Then $\lambda I - A$ is not onto but

$$\overline{\operatorname{ran}(\lambda I - A)} = \mathcal{H}$$

For any $x \in \mathcal{H}$, we have

$$\begin{aligned}
\|(\lambda I - A)x\|^2 &= \langle (\lambda I - A)x, (\lambda I - A)x\rangle \\
&= \langle (aI - A)x, (aI - A)x\rangle + \langle ibx, (aI - A)x\rangle + \langle (aI - A)x, ibx\rangle + \langle ibx, ibx\rangle \\
&= \|(aI - A)\|^2 + b^2 \|x\|^2 + ib(\langle (aI - A), x\rangle - \langle x, (aI - A)\rangle) \\
&= \|(aI - A)\|^2 + b^2 \|x\|^2 + ib(\langle (aI - A), x\rangle - \langle (aI - A), x\rangle) && \text{Self-Adjointness of } (a \\
&= \|(aI - A)\|^2 + b^2 \|x\|^2 \\
&\geq b^2 \|x\|^2
\end{aligned}$$

Therefore,

$$\|(\lambda I - A)x\| \geq b^2 \|x\|$$

for all $x \in \mathcal{H}$. Since $b \neq 0$, then we proved last quarter that $\|(\lambda I - A)x\|$ having a lower bound is equivalent to

$$\operatorname{ran}(\lambda I - A) \text{ is closed}$$

But this contradicts what we assumed before:

$$\operatorname{ran}(\lambda I - A) = \overline{\operatorname{ran}(\lambda I - A)} = \mathcal{H} \neq \operatorname{ran}(\lambda I - A)$$

■

Corollary 13.19.1. *If A is self-adjoint, then eigenvectors corresponding to different eigenvalues are orthogonal.*

Proof. Let λ_1 and λ_2 be two eigenvalues of A with $\lambda_1 \neq \lambda_2$. Let x_1 and x_2 be eigenvectors associated with λ_1 and λ_2 respectively. Then $\lambda_1, \lambda 2$ are real and:

$$\langle Ax_1, x_2\rangle = \langle x_1, Ax_2\rangle \implies \langle \lambda_1 x_1, x_2\rangle = \langle x_1, \lambda_2 x_2\rangle \implies \lambda_1\langle x_1, x_2\rangle = \lambda_2\langle x_1, x_2\rangle \implies \langle x_1, x_2\rangle = 0$$

■

14 Compact Operators in Banach Space

Definition 14.1. *Let X and Y be Banach spaces. A linear operator $T : X \to Y$ is compact if $T(B_1(0))$ is precompact in Y (or $\overline{T(B_1(0))}$ is compact in Y), where $B_1(0)$ is the open unit ball surrounding the origin.*

Remark 14.2. *$T : X \to Y$ is compact $\iff$ $T(B)$ is precompact in Y whenever B is a bounded subset of X $\iff$ for every bounded sequence $\{x_n\}$ in X, $\{T(x_n)\}$ has a convergent subsequence in Y.*

Definition 14.3. *$K(X,Y)$ is the space of all compact operators from $X \to Y$.*

Properties:

- $K(X,Y) \subset B(X,Y)$

  Proof. Consider for every $T \in K(X,Y)$,

$$\|T\| = \sup_{\|x\|\le 1} \|Tx\| = \sup_{y \in T(B_1)} \|y\| \le \max_{y\in\overline{T(B_1)}} \|y\| = M < +\infty$$

  because $\overline{T(B_1)}$ is compact. Therefore, $T \in B(X,Y)$. ■

- $K(X,Y)$ is a linear subspace of $B(X,Y)$.

  Proof. For all $T, S \in K(X,Y)$, $\alpha, \beta \in \mathbb{C}$ then we want to show $\alpha T + \beta S \in K(X,Y)$. Indeed for any bounded sequence (x_n) in X:

  - Since T is compact, $(T(x_n))$ has a convergent subsequence $T(x_{n_k})$ in Y.
  - Since S is compact, $(T(x_{n_k}))$ has a convergent subsequence $(T(x_{n_{k_\ell}}))$ in Y.

  Therefore, $((\alpha T + \beta s)(x_n))$ gas a convergent subsequence $((\alpha T + \beta s)(x_{n_{k_\ell}}))$ in Y. $\implies$ $\alpha T + \beta S \in K(X,Y)$ ■

- $K(X,Y)$ is closed in $B(X,Y)$. That is, if $T_n \in K(X,Y), \|T_n - T\| \to 0$ for some $T \in B(X,Y)$, then $T \in K(X,Y)$.

  Proof. For the unit ball B_1 of X, since T_n is compact, then $\overline{T_n(B_1)}$ is compact in Y. Thus, $\overline{T_n(B_1)}$ has a finite $\frac{\epsilon}{2}$-net $\{y_1, y_2, \ldots, y_m\}$. Then

$$\overline{T_n(B_1)} \subset \bigcup_{i=1}^{m} B_\epsilon(y_i)$$

  So $\overline{T_n(B_1)}$ has a finite ϵ-net. Therefore, $\overline{T_n(B_1)}$ is compact and $T \in K(X,Y)$ ■

- If $T \in B(X,Y)$ has finite dimensional range, then $T \in K(X,Y)$

  Proof. For every bounded sequence (x_n) in X, $(T(x_n))$ is a bounded sequence in a finite dimensional space. Thus (Tx_n) has a convergent subsequence. Therefore, T is compact. ■

- Suppose $T \in B(X,Y), S \in B(X,Y)$, if either T or S is compact, then $S \circ T \in K(X,Z)$.

Recall If $x_n \rightharpoonup x$ in $\mathcal{H}$ and $A \in B(\mathcal{H})$, then $Ax_n \rightharpoonup Ax$.

Proof. For all $z \in \mathcal{H}$, since $x_n \rightharpoonup x$, then

$$\langle Ax_n, z\rangle = \langle x_n, A^* z\rangle \to \langle x, A^* z\rangle = \langle Ax, z\rangle$$

■

Theorem 14.4. *Suppose $A \in B(\mathcal{H})$. Then A is compact if and only if A maps every weakly convergent subsequence into a strongly convergent sequence. That is,*

$$x_n \rightharpoonup x \in \mathcal{H} \implies Ax_n \to Ax$$

provided A is compact.

Proof. ($\Leftarrow$) For any bounded sequence (x_n) in $\mathcal{H}$, it has a weakly convergent subsequence (x_{n_k}) such that $x_{n_k} \rightharpoonup x$ for some $x \in \mathcal{H}$. By assumption, $Ax_{n_k} \to Ax$. Therefore, (Ax_n) has a strongly convergent subsequence (Ax_{n_k}). Therefore, A is compact.

($\Rightarrow$) Suppose A is compact and $x_n \rightharpoonup x$ in $\mathcal{H}$. Let $y_n = x_n - x$, then $y_n \rightharpoonup 0$ in $\mathcal{H}$. We want to show that $Ay_n \to 0 \implies \|Ay_n\| \to 0$. Since the norm is always positive, we only need to demonstrate the sequence lim sup goes to zero. That is, let

$$\sigma := \limsup_{n\to\infty} \|Ay_n\| \geq 0$$

Then there exists a subsequence (Ay_{n_k}) such that $\|Ay_{n_k}\| \to \sigma$. Since $y_{n_k} \rightharpoonup 0$, and $A \in B(\mathcal{H})$, we have $Ay_{n_k} \rightharpoonup A \cdot 0 = 0$. Also $y_{n_k} \rightharpoonup 0$ implies (y_{n_k}) is bounded. Since A is compact, (Ay_{n_k}) has a convergent subsequence $(y_{n_{k_\ell}})$

$$Ay_{n_{k_\ell}} \rightharpoonup z \text{ as } \ell \to \infty$$

But $y_{n_{k_\ell}} \rightharpoonup$ so $Ay_{n_{k_\ell}} \rightharpoonup 0$. Therefore, $z = 0$. Moreover, $\left\|Ay_{n_{k_\ell}}\right\| \to 0$ which implies $\sigma = 0$. Therefore,

$$\lim_{n\to\infty} \|Ay_n\| = 0$$

and $Ay_n \to 0$ as desired. ■

14.1 Compact Operators and their Spectrum

Proposition 14.5. *Let A be a compact operator on a Hilbert space $\mathcal{H}$. Then for all $\lambda \in \sigma_p(A) \setminus \{0\}$ has finite multiplicity. That is, when $E_\lambda = Ker(\lambda I - A)$ for some $\lambda \neq 0$, then $dim(E_\lambda) < \infty$.*

Proof. Assume $\dim(E_\lambda) = \infty$. Then it has an orthonormal eigenvectors $\{e_n\}_{n=1}^\infty$. Since $\{e_n\}_{n=1}^\infty$ is bounded and A is compact, the sequence

$$(Ae_n)_{n=1}^\infty = (\lambda e_n)_{n=1}^\infty$$

has a convergent subsequence. This is impossible because

$$\|\lambda e_n - \lambda e_m\| = |\lambda|\sqrt{2}\delta_{mn}$$

In particular, if $1 \in \sigma_p(A)$, then $\dim(\text{Ker}(I - A)) < \infty$. ■

Theorem 14.6 (Fredholm Alternative)**.** *Let $K : \mathcal{H} \to \mathcal{H}$ be a compact operator on a Hilbert space $\mathcal{H}$. Let $T = I - K$, so $T^* = I - K^*$. Then*

1. *$dim(Ker(T)) = dim(Ker(T^*) < \infty$*
2. *$range(T)$ is closed.*

3. *range*$(T) = Ker(T^*)^\perp$

4. *Ker*$(T) = \{0\} \iff$ *range*$(T) = \mathcal{H}$. *That is, T is one-to-one if and only if T is onto.*

Remark 14.7. *For all $\lambda \in \sigma(K) \setminus \{0\}$, $\lambda I - K$ is also compact. So $\lambda I - K$ is one-to-one if and only if $\lambda I - K$ is onto. $\implies \lambda \notin \sigma_c(K) \cup \sigma_r(K) \implies \lambda \in \sigma_p(K) \setminus \{0\}$. Moreover, when K is compact, $\sigma(K) \setminus \{0\} = \sigma_p(K) \setminus \{0\}$.*

Proposition 14.8. *When dim$(\mathcal{H}) = \infty$, a compact operator $K : \mathcal{H} \to \mathcal{H}$ can not have an inverse. That is, $0 \in \sigma(K)$.*

Proof. Let $\{e_n\}_n$ is an orthonormal basis of $\mathcal{H}$. Assume K has an inverse. Then $\{K^{-1}e_n\}$ is a bounded sequence. Since K is compact, then

$$\{e_n\} = \{K(K^{-1}e_n)\}$$

has a convergent subsequence. This contradicts proposition 8.5. Therefore, K has no inverse. ■

Proposition 14.9. *Let K be a compact and self-adjoint operator on a Hilbert space. Then*

$$\sigma(K) \setminus \{0\} = \sigma_p(K) \setminus \{0\} \subset [-\|K\|, \|K\|]$$

Proposition 14.10. *Any countable infinite set of nonzero eigenvalues $\{\lambda_n\}$ of a compact, self-adjoint operator K on $\mathcal{H}$ has zero as an accumulation point, and no other accumulation points.*

Proof. Let $\{\lambda_n\}_{n=1}^\infty$ be any non-zero eigenvalues of K, and $\{e_n\}$ be the sequence of associated eigenvectors with

$$Ke_n = \lambda_n e_n.$$

Since $\{\lambda_n\}_{n=1}^\infty \subset [-\|K\|, \|K\|] \implies \{\lambda_n\}_{n=1}^\infty$ has a convergent subsequence $\{\lambda_{n_k}\} \to \lambda \in [-\|K\|, \|K\|]$.

Claim: $\lambda = 0$
Assume otherwise. Let $f_{n_k} = \frac{1}{\lambda_{n_k}} e_{n_k}$. Then,

$$\|f_{n_k}\| = \frac{1}{|\lambda_{n_k}|} \to \frac{1}{\lambda}$$

which implies $\{f_{n_k}\}$ is bounded. Since K is compact,

$$\{Kf_{n_k}\} = \{e_{n_k}\}$$

has a convergent subsequence. This is impossible by Proposition 8.5. Therefore, $\lambda = 0$. ■

From this proposition, we see there are three possible cases when it comes to eigenvalues:

1. $\sigma(K) = \{0\}$ (Finite)

2. $\sigma(K) = \{\lambda_1, \ldots, \lambda_n\}$ (Finite)

3. $\sigma(K) = \{0, \lambda_1, \lambda_2, \ldots\}$ (Infinite with accumulation point 0)

Lemma 14.11. *If K is a compact, self-adjoint operator on $\mathcal{H}$, then at least one of $\pm\|K\|$ is an eigenvalue of K.*

Proof. Recall that if K is bounded, self-adjoint then

$$\|K\| = \sup_{\|x\|=1} |\langle x, Kx \rangle|$$

We want to prove that there exists $x_0 \in \mathcal{H}, \|x\| = 1$ such that

$$\sup_{\|x\|=1} |\langle x, Kx\rangle| = |\langle x_0, Kx_0\rangle|$$

Moreover, $Kx_0 = \lambda x_0$ for some $|\lambda| = \|K\|$. We may assume that

$$\|K\| = \sup_{\|x\|=1} |\langle x, Kx\rangle| = \sup_{\|x\|=1} \langle x, Kx\rangle = \lambda$$

Otherwise, we consider $-K$ instead.
Taking the maximizing sequence $\{x_n\}$ with $\|x_n\| = 1$ such that

$$\langle x_n, Kx_n\rangle \to \lambda$$

Since $\{x_n\}$ is bounded, it has a weakly convergent subsequence, still denoted by x_n, such that

$$x_n \rightharpoonup x_0$$

So $\|x_0\| \leq \liminf_{n\to\infty} \|x_n\| = 1$ Since K is compact, $x_n \rightharpoonup x_0$, then we must have $Kx_n \to Kx_0$. So,

$$\langle x_n, Kx_n\rangle \to \langle x_0, Kx_0\rangle$$

Therefore, $\langle x_0, Kx_0\rangle = \lambda$. Also,

$$\lambda = \langle x_0, Kx_0\rangle \leq \|K\| \|x_0\|^2 \leq \|K\| = \lambda$$

Claim: $Kx_0 = \lambda x_0$
Indeed,

$$\begin{aligned}\|Kx_n - \lambda x_n\|^2 &= \|Kx_n\|^2 + |\lambda|^2 \|x_n\|^2 + 2\mathrm{Re}\langle Kx_n, -\lambda x_n\rangle \\ &\leq \|K\|^2 \|x_n\|^2 + |\lambda|^2 \|x_n\|^2 - 2\lambda\langle Kx_n, x_n\rangle\end{aligned}$$

Letting $n \to \infty$

$$\|Kx_0 - \lambda x_0\|^2 \leq \|K\|^2 + |\lambda|^2 - 2\lambda \cdot \lambda = 0$$

Therefore, $Kx_0 = \lambda x_0 \implies \lambda = \|K\|$ is an eigenvalue of K. ■

Corollary 14.11.1. *$\sigma_p(K)$ is either finite or countably infinite.*

Proof. Since

$$\sigma_p(K) = \bigcup_n [-\|K\|, \|K\|] \setminus (\frac{-1}{n}, \frac{1}{n}) \cap \sigma_p(K)$$

since 0 is the only allowable accumulation point, then we know: $[-\|K\|, \|K\|] \setminus (\frac{-1}{n}, \frac{1}{n}) \cap \sigma_p(K)$ is finite. Therefore, the union of finite sets is at mount countable. ■

Now, for any $\lambda \in \sigma_p(K) \setminus \{0\}, m_\lambda = \dim(E_\lambda) < \infty$ where $E_\lambda = \mathrm{Ker}(\lambda I - K)$ is the eigenspace. Let $\{e_i^{(\lambda)}\}_{i=1}^{m_\lambda}$ be an orthonormal basis of E_λ. If $0 \in \sigma_p(K)$, then let $\{e_i^{(0)}\}$ be an orthonormal basis of $\mathrm{Ker}(K)$. Notice, $\{e_i^{(0)}\}$ is not necessarily countable. Let $U = \{e_i\} = \bigcup_{\lambda\in\sigma_p(A)} \{e_i^{(\lambda)}\}$

Claim: $U = \{e_i\}$ is an orthonormal basis of $\mathcal{H}$.

Proof. We prove the equivalent claim $U^\perp = \{0\}$. Assume otherwise. Then we notice that $U^\perp$ is a closed linear subspace of $\mathcal{H}$, so $U^\perp$ itself is also Hilbert. Also $x \in U^\perp, Kx \in U^\perp$ since

$$\langle Kx, e_i\rangle = \langle x, Ke_i\rangle = \langle x, \lambda_i e_i\rangle = \lambda_i\langle x, e_i\rangle = 0$$

Define $\tilde{K} = K|_{U^\perp} : U^\perp \to U^\perp$. By the previous lemma, $\tilde{x}_0 \in U^\perp$ such that $\|\tilde{x}_0\| = 1$ and $\tilde{K}\tilde{x}_0 = \tilde{\lambda}\tilde{x}_0$ for some $|\tilde{\lambda}| = \left\|\tilde{K}\right\|$. So $K\tilde{x}_0 = \tilde{\lambda}x_0 \implies \tilde{\lambda} \in \sigma_p(K)$ and $\tilde{x}_0 \in E_{\tilde{\lambda}}$. But this contradicts the fact that $\tilde{x}_0 \in U^\perp$ and $\|\tilde{x}_0\| = 1$. Therefore, $U^\perp = \{0\}$ and $U = \{e_i\}$ is an orthonormal basis of $\mathcal{H}$. ■

Now, for every $x \in \mathcal{H}$, we can write it as

$$x = \sum_{i=1}^{\infty} \langle e_i, x \rangle e_i.$$

In particular,

$$\begin{aligned}
Kx &= \sum_{i=1}^{\infty} \langle e_i, Kx \rangle e_i \\
&= \sum_{i=1}^{\infty} \langle Ke_i, x \rangle e_i \\
&= \sum_{i=1}^{\infty} \langle \lambda_i e_i, x \rangle e_i \\
&= \sum_{i=1}^{\infty} \overline{\lambda_i} \langle e_i, x \rangle e_i \\
&= \sum_{i=1}^{\infty} \lambda_i \langle e_i, x \rangle e_i \\
&= \sum_{\lambda_i \neq 0} \lambda_i \langle e_i, x \rangle e_i \\
&= \sum_{\lambda \in \sigma_p(K) \setminus \{0\}} \lambda \sum_{i=1}^{m(\lambda)} \langle e_i^{(\lambda)}, x \rangle e_i^{(\lambda)} \\
&= \sum_{\lambda \in \sigma_p(K) \setminus \{0\}} \lambda P_\lambda(x)
\end{aligned}$$

Where

$$\begin{aligned}
P_\lambda : \mathcal{H} &\to E_\lambda \\
x &\to \sum_{i=1}^{m(\lambda)} \langle e_i^{(\lambda)}, x \rangle e_i^{(\lambda)}
\end{aligned}$$

is the orthogonal projection. Therefore,

$$K = \sum_{\lambda \in \sigma_p(K) \setminus \{0\}} \lambda P_\lambda$$

Theorem 14.12 (Spectral Theorem for Compact, Self-Adjoint Operators). *Let K be a compact, self-adjoint operator on a Hilbert space $\mathcal{H}$. Then*

1. *K has at most countable nonzero eigenvalues $\{\lambda_k\}$ of real numbers with 0 a the only possible accumulation point.*
2. *There is an orthonormal basis $\{e_i\}$ of $\mathcal{H}$ consisting of eigenvectors of K such that for all $x = \sum_{i=1}^{\infty} \langle e_i, x \rangle e_i$, then*

$$K(x) = \sum_{i=1}^{\infty} \lambda_i \langle e_i, x \rangle e_i$$

3. *Further,*
$$K = \sum_{\lambda_k \in \sigma_p(K)\setminus\{0\}} \lambda_k P_{\lambda_k}$$
where $P_k : \mathcal{H} \to E_{\lambda_k}$ is the orthogonal projection.

Remark 14.13. *One may order the eigenvalues of K based on its absolute value:*
$$|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n| \geq |\lambda_{n+1}| \geq \ldots \to 0$$
with the convention that each eigenvalue is assigned as many as its multiplicity.

Claim: $K = \sum_{i=1}^{\infty} \lambda_i \langle e_i, \cdot \rangle e_i$ where the sum converges in the operator norm.

Proof. Indeed, for every $x \in \mathcal{H}$,
$$\begin{aligned}
\left\| Kx - \sum_{i=1}^{n} \lambda_i \langle e_i, x\rangle e_i \right\| &= \left\| \sum_{i=n+1}^{\infty} \lambda_i \langle e_i, x\rangle e_i \right\| \\
&= \left(\sum_{i=n+1}^{\infty} \lambda_i |\langle e_i, x\rangle|^2 \right)^{\frac{1}{2}} && \text{Parseval's Identity} \\
&\leq |\lambda_{n+1}| \left(\sum_{i=n+1}^{\infty} |\langle e_i, x\rangle|^2 \right)^{\frac{1}{2}} \\
&\leq |\lambda_{n+1} \, \|x\|
\end{aligned}$$

So,
$$\left\| K - \sum_{i=1}^{n} \lambda_i \langle e_i, \cdot\rangle e_i \right\| \leq |\lambda_{n+1}| \to 0$$

Therefore,
$$K = \sum_{i=1}^{\infty} \lambda_i \langle e_i, \cdot\rangle e_i$$

■

Remark 14.14. *One may calculate the λ_i's as follows:*
$$|\lambda_1| = \|K\| = \sup_{\|x\|=1} |\langle x, Kx\rangle|$$

- *Positive if*
$$\sup_{\|x\|=1} |\langle x, Kx\rangle| = \sup_{\|x\|=1} \langle x, Kx\rangle$$
- *Otherwise, negative.*

Further,
$$|\lambda_2| = \|K - \lambda_i \langle e_i, \cdot\rangle e_i\| = \sup_{\|x\|=1, x \perp span\{e_1\}} \langle x, Kx\rangle$$
$$\vdots$$
$$|\lambda_n| = \left\| K - \sum_{i=1}^{n-1} \lambda_i \langle e_i, \cdot\rangle e_i \right\| = \sup_{\|x\|=1, x \perp span\{e_1, \ldots, e_{n-1}\}} \langle x, Kx\rangle$$

Remark 14.15. *Compact, self-adjoint operator is diagonalizable*

$$K = \sum_{i=1}^{\infty} \lambda_i \langle e_i, \cdot \rangle e_i$$

For all x,

$$\begin{aligned} Kx &= \sum_{i=1}^{\infty} \lambda_i \langle e_i, x \rangle e_i \\ &= \begin{pmatrix} x_1 & x_2 & \dots \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \dots \\ 0 & \lambda_2 & \vdots \\ \vdots & \dots & \ddots \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \vdots \end{pmatrix} \end{aligned}$$

14.2 Fredholm Alternative

Recall the Theorem:

Theorem 14.16 (Fredholm Alternative)**.** *Let $K : \mathcal{H} \to \mathcal{H}$ be a compact operator on a Hilbert space $\mathcal{H}$. Let $T = I - K$, so $T^* = I - K^*$. Then*

1. *dim(Ker(T)) = dim(Ker(T*) < ∞* — $\dim(\operatorname{Ker}(T)) = \dim(\operatorname{Ker}(T^*) < \infty$
2. *range(T) is closed.*
3. *range(T) = Ker(T^*)$^\perp$*
4. *Ker(T) = $\{0\}$ $\iff$ range(T) = $\mathcal{H}$. That is, T is one-to-one if and only if T is onto.*

For every compact operator K on $\mathcal{H}$, we can partition the space so that K belongs to exactly one of the following alternatives:

1. $$K \in \mathcal{F} = \{K \text{ compact } : \mathrm{Ker}(I - K) = \{0\}\}$$

2. $$K \in \mathcal{F}^c = \{K \text{ compact } : \mathrm{Ker}(I - K) \neq \{0\}\}$$

What properties can we say of these two cases:

1. *Case:* $K \in \mathcal{F}$, therefore in this case, $\mathrm{Ker}(I-K) = \{0\}$ and $\mathrm{Range}(I-K) = \mathcal{H}$ so $I-K$ is a one-to-one and onto. Therefore,
   (a) $1 \in \rho(K)$.
   (b) Further, for all $y \in \mathcal{H}$, the equation
   $$x - Kx = y$$
   has a unique solution.
   (c) $K \in \mathcal{F} \iff \mathrm{range}(I-K) = \mathcal{H} \iff (\mathrm{Ker}(I-K^*))^\perp = \mathcal{H} \implies \mathrm{Ker}(I-K^*) = \{0\} \implies K^* \in \mathcal{F}$ Therefore, $K \in \mathcal{F} \iff K^* \in \mathcal{F}$.
2. *Case:* $K \in \mathcal{F}^c$, In this case, $\mathrm{Ker}(I-K) \neq \{0\}$. Then
   (a) $1 \in \sigma(K) \setminus \{0\} = \sigma_p(K) \setminus \{0\} \implies \dim(E_1) < \infty$

(b) The equation

$$x - Kx = 0$$

has a nontrivial solution. This nontrivial solution forms a finite dimensional subspace of $\mathcal{H} \implies E_1$

(c) $K \in \mathcal{F} \implies K^* \in \mathcal{F}^c$

(d) For each $y \in \mathcal{H}$, the inhomogenous equation

$$x - Kx = y$$

has a non-unique solution if and only if

$$y \in \operatorname{ran}(I - K) = (\operatorname{Ker}(I - K^*))^{\perp} \neq \mathcal{H}$$

if and only if $y \perp z$ for every solution z of $z - K^* z = 0$.

15 4 Modes of Convergence

There are 4 modes of convergence. They are:

1. Almost everywhere (a.e.) convergence
2. Almost uniform (a.u.) convergence
3. Convergence in measure
4. L^p convergence (mean convergence)

Definition 15.1. *Given $(X, \mathscr{A}, \mu)$ measure space, $f_n : D \to \mathbb{R}$, on a $\mathscr{A}$-measurable set $D \in \mathscr{A}$, $f : D \to \mathbb{R}$ a $\mathscr{A}$-measurable function, then $f_n \to f$ almost everywhere provided there exists $A \subset D, A \in \mathscr{A}, \mu(A) = 0$ such that*

$$x \in D \setminus A \implies f_n(x) \to f(x)$$
$$\iff \mu(\{x \in D : f_n(x) \not\to f(x)\}) = 0$$

Recall the definitions:

$$\limsup E_n = \bigcap_{k=1}^{\infty} \bigcup_{n \geq k} E_n = \{x \in E_n \text{ for infinitely many } n\}$$

$$\liminf E_n = \bigcup_{k=1}^{\infty} \bigcap_{n \geq k} E_n = \{x \in E_n \text{ for all but finitely many } E_n\}$$

Remark 15.2. *Convergence Almost Everywhere is equivalent to for a given ϵ*

$$\iff \mu(\{x \in D : |f_n(x) - f(x)| \geq \epsilon \textit{ for infinitely many } n\}) = 0$$

To better construct this, we define for a given $\epsilon > 0$,

$$E_n^{\epsilon} = \{x \in D : |f_n(x) - f(x)| \geq \epsilon\}$$

Therefore, the set:

$$E^{\epsilon} = \mu(\{x \in D : |f_n(x) - f(x)| \geq \epsilon \textit{ for infinitely many } n\})$$

must be equal to the lim sup of E_n^{ϵ}:

$$E^{\epsilon} = \bigcap_{k=1}^{\infty} \bigcup_{n \geq k} E_n^{\epsilon}$$

Therefore, convergence almost everywhere requires:

$$\mu(E^\epsilon) = \mu(\limsup_{n\to\infty} E_n^\epsilon) = \mu\left(\bigcap_{k=1}^{\infty}\bigcup_{n\geq k} E_n^\epsilon\right) = 0$$

Definition 15.3. *Given $(X, \mathscr{A}, \mu)$ a measure space, let $f_n : D \to \mathbb{R}$, a $\mathscr{A}$-measurable function defined on $D \in \mathscr{A}$, and $f : D \to \mathbb{R}$ another $\mathscr{A}$-measurable function, then $f_n \to f$ almost uniformly on D provided for all $\epsilon > 0$, $\exists A_\epsilon \in \mathscr{A}, A_\epsilon \subset D$ such that*

$$\mu(A_\epsilon) < \epsilon \textit{ and } f_n \to f \textit{ uniformly on } D \setminus A_\epsilon = A_\epsilon^c$$

Definition 15.4. *Given $(X, \mathscr{A}, \mu)$ a measure space, let $f_n : D \to \mathbb{R}$, a $\mathscr{A}$-measurable function defined on $D \in \mathscr{A}$, and $f : D \to \mathbb{R}$ another $\mathscr{A}$-measurable function, then $f_n \underset{\mu}{\to} f$ in measure on D provided for all $\epsilon > 0$,*

$$\lim_{n\to\infty} \mu(\{x \in D : |f_n(x) - f(x)| \geq \epsilon\}) = 0$$

Definition 15.5. *Given $(X, \mathscr{A}, \mu)$ a measure space, let $f_n : D \to \mathbb{R}$, a $\mathscr{A}$-measurable function defined on $D \in \mathscr{A}$, and $f : D \to \mathbb{R}$ another $\mathscr{A}$-measurable function, then $f_n \to f$ in L^p on D provided*

$$\lim_{n\to\infty} \int_D |f_n - f|^p \, d\mu = 0$$

with $\infty > p \geq 1$.

Remark 15.6. *It's important to see the immediate relationships between the different modes of convergence. Specifically, we want to show:*

$$\begin{tikzcd}
a.e. & a.u \arrow[l] \arrow[d] \\
L^p \arrow[r] & \mu
\end{tikzcd}$$

where the arrows indicate implication.

Lemma 15.7. *Almost uniform convergence implies almost every convergence.*

Proof. Given $(X, \mathscr{A}, \mu)$ a measure space, let $f_n : D \to \mathbb{R}$, a $\mathscr{A}$-measurable function defined on $D \in \mathscr{A}$, and $f : D \to \mathbb{R}$ another $\mathscr{A}$-measurable function, suppose $f_n \to f$ almost uniformly.
Then for every $\epsilon > 0$, $\exists A_\epsilon \in \mathscr{A}, A_\epsilon \subset D$ such that

$$\mu(A_\epsilon) < \epsilon \text{ and } f_n \to f \text{ uniformly on } D \setminus A_\epsilon = A_\epsilon^c$$

Pick $\epsilon = \frac{1}{n}$. Then we see the sequence $\{A_{\frac{1}{n}}\}_n$, with $\mu(A_{\frac{1}{n}}) < \frac{1}{n}$, and can define the set:

$$A := \bigcap_{n=1}^{\infty} \implies \mu(A) \leq \mu(A_{\frac{1}{n}}) < \frac{1}{n}$$

Now, if $x \in D \setminus A \implies \mu(A) = 0 \implies x \in A_{\frac{1}{N}}$ for some $N \implies f_n(x) \to f(x) \implies$ therefore f_n converges to f almost everywhere. ∎

Lemma 15.8. *Almost uniform convergence implies convergence in measure.*

Proof. Given $(X, \mathscr{A}, \mu)$ a measure space, let $f_n : D \to \mathbb{R}$, a $\mathscr{A}$-measurable function defined on $D \in \mathscr{A}$, and $f : D \to \mathbb{R}$ another $\mathscr{A}$-measurable function, suppose $f_n \to f$ almost uniformly.
Then for every $\epsilon > 0$, $\exists A_\epsilon \in \mathscr{A}, A_\epsilon \subset D$ such that

$$\mu(A_\epsilon) < \epsilon \text{ and } f_n \to f \text{ uniformly on } D \setminus A_\epsilon = A_\epsilon^c$$

Given $\delta > 0$, $\exists N$ such that

$$|f_n(x) - f(x)| < \delta$$

for all $n > N, x \in D \setminus A_\epsilon$. Therefore, the set:

$$\{x \in D : |f_n(x) - f(x)| \geq \delta\} \subset D \setminus A_\epsilon \text{ for } n > N$$

Taking the measures,

$$\mu(\{x \in D : |f_n(x) - f(x)| \geq \delta\}) \leq \mu(A_\epsilon) < \epsilon$$

Since this is true for arbitrary ϵ, n we conclude

$$\lim_{n\to\infty} \mu(\{x \in D : |f_n(x) - f(x)| \geq \delta\}) = 0$$

Therefore, we have convergence in measure! ∎

Lemma 15.9 (Chebyshev's Inequality)**.** *Suppose $f \in L^p(D)$, then*

$$\mu(\{x \in D : |f(x)| \geq \epsilon\}) \leq \frac{1}{\epsilon^p} \int_D |f|^p \, d\mu$$

Proof.

$$\begin{aligned}
\int_D |f|^p \, d\mu &\geq \int_{\{x \in D : |f(x)| \geq \epsilon\}} |f|^p \, d\mu \\
&\geq \int_{\{x \in D : |f(x)| \geq \epsilon\}} \epsilon^p \, d\mu \\
&= \epsilon^p \mu(\{x \in D : |f(x)| \geq \epsilon\})
\end{aligned}$$

∎

Lemma 15.10. *L^p convergence implies convergence in measure.*

Proof. Suppose $f_n \to f$ in L^p, then

$$\lim_{n\to\infty} \int_D |f_n - f|^p d\mu = 0$$

By Chebyshev's Inequality,

$$\mu(\{x \in D : |f_n(x) - f(x)| \geq \epsilon\}) \leq \frac{1}{\epsilon^p} \int_D |f_n - f|^p \, d\mu$$

$$\implies \lim_{n\to\infty} \mu(\{x \in D : |f_n(x) - f(x)| \geq \epsilon\}) \leq \frac{1}{\epsilon^p} \int_D |f_n - f|^p \, d\mu \to 0$$

Therefore, we have convergence in measure! ∎

Now, let's consider implication relationships that do not exist between these modes of convergence.

Example 15.11. *Consider the space $(X, \mathscr{B}(X), \lambda)$, and define the sequence*

$$\{f_n = \mathcal{X}_{[n,n+1)}\}_{n=1}^{\infty}$$

Notice $f_n \to 0$ almost everywhere BUT

$$\lim_{n\to\infty} \mu(\{x \in \mathbb{R} : |f_n| > 1/2\}) = 0 \implies f_n \underset{\mu}{\not\to} 0$$

Also,

$$f_n \not\to 0 \text{ a.e.}$$

Lastly, since

$$\int_{\mathbb{R}} |f_n|^p \; d\mu = 1 \implies f_n \not\to 0 \text{ in } L^p$$

Example 15.12. *Consider the space $(X, \mathscr{B}(X), \lambda)$, and define the sequence Define the function the sequence of functions $\{f_n\}_n$ by*

$$f_n(x) = \mathcal{X}_{A_n}(x)$$

where

$$A_n := [j2^{-k}, (j+1)2^{-k}]$$

where we decompose $n = 2^k + j$. Clearly, as $n \to \infty$, $k \approx \log_2(n) \to \infty$. Notice,

$$\int_{\mathbb{R}} f_n \; d\lambda = \int_{\mathbb{R}} \mathcal{X}_{A_n} \; d\lambda = \mu(A_n) = (j+1)2^{-k} - j2^{-k} = 2^{-k} \to 0 \text{ as } n \to \infty$$

However, for every x, we can define a subsequence n_k such that

$$f_{n_i}(x) \to 1 \text{ as } i \to \infty$$

since we can always find an integer $n = 2^k + j$ such that $j2^{-k} \leq x \leq (j+1)2^{-k}$. On the other hand, we can just as well as find a subsequence such that

$$f_{n_\ell}(x) \to 0 \text{ as } i \to \infty$$

Because these two subsequence fail to converge to the same functional value, we conclude $f_n(x) \not\to f(x)$ a.e. Therefore, since we have L^p convergence, then we know f_n converges in measure. Also, since the sequence fails to converge almost everywhere, then it must also fail to converge uniformly.

Example 15.13. *Consider the space $(X, \mathscr{B}(X), \lambda)$, and define the sequence*

$$\{f_n = n\mathcal{X}_{[0,\frac{1}{n}]}\}_{n=1}^{\infty}$$

Therefore, we see:

$$\int_{\mathbb{R}} |f_n| d\lambda = 1 \implies f_n \not\to 0 \text{ in } L^1$$

$$f_n \to 0 \text{ in } \lambda$$

$$f_n \to 0 \text{ a.u.}$$

$$f_n \to 0 \text{ a.e.}$$

As we will see, we can consider more relationships between the modes of convergence, but these implications require stronger assumptions. But before we discuss these, a review of some measure theory results:

Lemma 15.14 (Borel-Cantelli)**.** *Suppose we are given $(X, \mathscr{A}, \mu)$ a measure space and $\{E_n\}_{n=1}^\infty \subset \mathscr{A}$. Suppose*

$$\sum_{n=1}^{\infty} \mu(E_n) < \infty$$

then

$$\mu(\limsup_{n\to\infty} E_n) = 0$$

Recall two facts about measure: Let $\{F_k\}_k \subset \mathscr{A}$ a set seqence

- (Increasing Sequence) If $F_k \subset F_{k+1}$ for all k , then $\lim_{k\to\infty} \mu(F_k) = \mu(\lim_{k\to\infty} F_k)$
- (Decreasing Sequence) If $F_{k+1} \subset F_k$ and $\mu(X) < \infty$, then $\lim_{k\to\infty} \mu(F_k) = \mu(\lim_{k\to\infty} F_k)$

Proof. All we have to do is unravel the definition of limsup:

$$\limsup_{n\to\infty} E_n = \bigcap_{k=1}^{\infty} \bigcup_{n\geq k} E_n$$

Define, $F_k := \bigcup_{n\geq k} E_n$. Since $\mu(F_1) < \infty$, then we see:

$$\begin{aligned}
\mu(\limsup_{n\to\infty} E_n) &= \mu\left(\bigcap_{k=1}^{\infty} \bigcup_{n\geq k} E_n\right) \\
&= \mu\left(\bigcap_{k=1}^{\infty} F_k\right) \\
&= \lim_{k\to\infty} \mu(F_k) && \text{Decreasing Sequence} \\
&= \lim_{k\to\infty} \mu(\bigcup_{n\geq k} E_n) \\
&\leq \lim_{k\to\infty} \sum_{n=k}^{\infty} \mu(E_n) = 0
\end{aligned}$$

∎

Corollary 15.14.1. *Let $(X, \mathscr{A}, \mu)$ be a measure space and let $f_n, f : D \to \mathbb{R}$ an $\mathscr{A}$-measurable function defined on $D \in \mathscr{A}$. Suppose the following conditions are true:*

1. *$\exists a_n > 0$ such that $\lim_{n\to\infty} a_n = 0$*
2. *$\sum_{n=1}^{\infty} \mu(\{x \in D : |f_n(x) - f(x)| \geq a_n\}) < \infty$*

Then $f_n \to f$ a.e. on D.

Proof. Given $\epsilon > 0$, define

$$E_n^\epsilon := \{x \in D : |f_n(x) - f(x)| \geq \epsilon\}$$

We need to show

$$\mu(\limsup_{n\to\infty} E_n^\epsilon) = 0$$

Since $\lim_{n\to\infty} a_n = 0$, then $\exists N$ such that $0 < a_n < \epsilon$ for all $n > N$. $\implies E_n^\epsilon \subset \underbrace{\{x \in D : |f_n(x) - f(x)| \geq a_n\}}_{F_n}$

Therefore, by Borel-Cantelli Lemma, we see by condition 2:

$$\sum_{n=1}^{\infty} \mu(F_n) < \infty \implies \mu(\limsup_{n\to\infty} F_n) = 0$$

Further, $\limsup_{n\to\infty} E_n^\epsilon \subset \limsup_{n\to\infty} F_n$ and therefore for all $\epsilon > 0$

$$\mu(\limsup_{n\to\infty} E_n^\epsilon) = 0$$

Which is equivalent to almost everywhere convergence. ■

Theorem 15.15 (Riesz)**.** *Let $(X, \mathscr{A}, \mu)$ be a measure space. Suppose $f_n, f : D \to \mathbb{R}$ an $\mathscr{A}$-measurable function defined on $D \in \mathscr{A}$. If $f_n \underset{\mu}{\to} f$, then $\exists$ a subsequence $f_{n_j} \to f$ a.e.*

Proof. This is actually a corollary to the previous corollary. If

$$\lim_{n\to\infty} \mu(\{x \in D : |f_n(x) - f(x)| \geq \epsilon\}) = 0$$

for all ϵ, then $n > N_\epsilon$ such that

$$\mu(\{x \in D : |f_n(x) - f(x)| \geq \epsilon\}) < \epsilon$$

Choose n_1 such that

$$\mu(\{x \in D : |f_{n_1}(x) - f(x)| \geq \frac{1}{2}\}) < \frac{1}{2}$$

Choose n_2 such that

$$\mu(\{x \in D : |f_{n_2}(x) - f(x)| \geq \frac{1}{2^2}\}) < \frac{1}{2^2}$$

$$\vdots$$

Choose n_j such that

$$\mu(\{x \in D : |f_{n_j}(x) - f(x)| \geq \frac{1}{2^j}\}) < \frac{1}{2^j}$$

Clearly, the sums are then:

$$\sum_{j=1}^{\infty} \mu(\{x \in D : |f_{n_j}(x) - f(x)| \geq \frac{1}{2^j}\}) \leq \sum_{j=1}^{\infty} \frac{1}{2^j} = 1$$

Therefore, by the previous corollary, we conclude the sequence $f_{n_j} \to f$ a.e. ■

Note 15.16. *As a result of Riesz' Theorem we can update our relation diagram to include this subsequence relationship:*

$$\begin{tikzcd} a.e. & a.u \arrow[l] \arrow[d] \\ L^p \arrow[r] & \mu \arrow[ul, dashed] \end{tikzcd}$$

where the dotted line indicates sub-sequential convergence.

15.1 Finite Measure Assumption

Theorem 15.17 (Lebesgue)**.** *Let $(X, \mathscr{A}, \mu)$ be a finite measure space. Suppose $f_n, f : D \to \mathbb{R}$ an $\mathscr{A}$-measurable function defined on $D \in \mathscr{A}$. If $f_n \to f$ a.e., then $f_n \underset{\mu}{\to} f$.*

Proof. Given $\epsilon > 0, E_n^\epsilon = \{x \in D : |f_n - f| \geq \epsilon\}$, then

$$\mu\left(\bigcap_{k=1}^{\infty} \underbrace{\bigcup_{n \geq k} E_n^\epsilon}_{F_k}\right) = 0$$

Again, F_k is a decreasing sequence, so

$$0 = \mu(\lim_{k\to\infty} F_k) = \lim_{k\to\infty} \mu(F_k)$$

$$\implies F_k \subset E_k^\epsilon \implies \lim_{k\to\infty} \mu(E_k^\epsilon) = 0$$

for all ϵ. This is equivalent to saying $f_n \underset{\mu}{\to} f$. ■

Note 15.18. *Updating our diagram, we see:*

$$\begin{tikzcd}
a.e. \arrow[dr, bend right, "\mu(X)<\infty"] & a.u \arrow[l] \arrow[d] \\
L^p \arrow[r] & \mu \arrow[ul, dashed, bend right]
\end{tikzcd}$$

Theorem 15.19 (Egorov)**.** *Let $(X, \mathscr{A}, \mu)$ be a finite measure space. Suppose $f_n, f : D \to \mathbb{R}$ an $\mathscr{A}$-measurable function defined on $D \in \mathscr{A}$. If $f_n \to f$ a.e., then $f_n \to f$ a.u.*

Theorem 15.20 (Generalized Egorov)**.** *Let $(X, \mathscr{A}, \mu)$ be a finite measure space. Suppose $f_n, f : D \to \mathbb{R}$ an $\mathscr{A}$-measurable function defined on $D \in \mathscr{A}$. Suppose the following conditions hold:*

1. *$f_n \to f$ a.e.*
2. *$\mu\left(\bigcup_{n=1}^{\infty}\{x \in X : |f_n(x) - f(x)| \geq \frac{1}{m}\}\right) < \infty$ for all $m \in \mathbb{N}$*

Then $f_n \to f$ a.u.

Proof. Given $\epsilon > 0$, we need to find $B \in \mathscr{A}, \mu(B) < \epsilon$ such that $f_n \to f$ uniformly on B^c. Define:

$$E := \bigcap_{k\geq 1} \bigcup_{n \geq k} \{x \in X : |f_n(x) - f(x)| \geq \frac{1}{m}\}$$

We want to show $\mu(E) = 0$ for all m. Now define

$$B_{k,m} := \bigcup_{n\geq k} \{x \in X : |f_n(x) - f(x)| \geq \frac{1}{m}\}$$

Notice $B_{k,m}$ is a decreasing sequence of sets in k AND $\mu(B_{1,m}) < \infty$ for all m by condition 2 of the generalized Egorov or the finitedness of the typical Egorov theorem. Therefore,

$$\mu(\lim_{k\to\infty} B_{k,m}) = \mu\left(\bigcap_{k\geq 1} B_{k,m}\right) = \mu(E) = 0 \implies \lim_{k\to\infty} \mu(B_{k,m}) = 0$$

Therefore, $\exists k_m$ such that

$$\mu(B_{k,m}) < \frac{\epsilon}{2^m}$$

Now let

$$B := \bigcup_{m\geq 1} B_{k_m,m}$$

To see that B satisfies our search, we check:

$$\mu(B) \leq \sum_{m\geq 1} \mu(B_{k_m,m}) < \sum_{m\geq 1} \frac{\epsilon}{2^m} = \epsilon$$

Second, we know $B^c = \bigcap_{m\geq 1} B^c_{k_m,m}$. So for a given $\delta > 0$, we let $M \in \mathbb{N}$ such that $\frac{1}{M} < \delta$. Then we see that for any $x \in B^c \implies x \in B^c_{k_M,M} \implies |f_n(x) - f(x)| < \frac{1}{M} < \delta$ for all $n > k_M$. Therefore, $f_n \rightrightarrows f$ at x! Therefore, $f_n \rightrightarrows f$ uniformly on B^c, completing the proof. ■

Note 15.21. *Updating our diagram to include this fact, we see:*

$$\begin{tikzcd}
a.e. \arrow[r, bend left, "\mu(X)<\infty"] \arrow[dr, bend right, "\mu(X)<\infty"] & a.u \arrow[l] \arrow[d] \\
L^p \arrow[r] & \mu \arrow[ul, dashed, bend right]
\end{tikzcd}$$

15.2 Dominated Sequence Assumption

Theorem 15.22 (Monotone Convergence)**.** *Given $(X, \mathscr{A}, \mu)$ be a measure space. Suppose $f_n, f : X \to [0, \infty)$ an $\mathscr{A}$-measurable function and suppose $f_n \uparrow$ (i.e. $f_{n+1} \geq f_n$). Then*

$$\lim_{n\to\infty} f_n = f \ a.e.$$

and

$$\int_X \lim_{n\to\infty} f_n \ d\mu = \lim_{n\to\infty} \int_X f_n d\mu$$

Proof. (Sketch)
Clearly, by monotonicity:

$$\int_X \lim_{n\to\infty} f_n \ d\mu \geq \lim_{n\to\infty} \int_X f_n d\mu$$

On the other hand, let $f := \lim_{n\to\infty} f_n$. Fix $0 \leq \phi(x) \leq f(x)$ a simple function ϕ. Define the set

$$A_n := \{x \in X : \exists\gamma \in (0,1) \text{ such that } f_n(x) \geq \gamma\phi(x)\}$$

Notice, $A_{n+1} \supset A_n \implies \mu(\lim A_n) = \lim \mu(A_n)$. Therefore,

$$\int_X f_n d\mu \geq \int_{A_n} \gamma\phi(x) d\mu \to \int_X \gamma\phi(x) d\mu \geq \int_X f d\mu = \int_X \lim_{n\to\infty} f_n d\mu$$

$$\implies \int_X \lim_{n\to\infty} f_n \ d\mu \leq \lim_{n\to\infty} \int_X f_n d\mu$$

■

Lemma 15.23 (Fatou). *Given $(X, \mathscr{A}, \mu)$ a measure space and suppose $f_n : X \to [0, \infty)$ an $\mathscr{A}$-measurable function. Then*

$$\int_X \liminf f_n d\mu \leq \liminf \int_X f_n d\mu$$

Proof. Recall ,

$$\liminf_{n\to\infty} f_n = \lim_{k\to\infty} \underbrace{\inf_{n\geq k} f_n}_{g_k}$$

Notice, $g_k \geq g_{k+1}$. Then

$$\begin{aligned} \int \liminf_{n\to\infty} f_n &= \int \lim_{k\to\infty} g_k \\ &= \lim_{k\to\infty} \int g_k && \text{Monotone Convergence Theorem} \\ &= \liminf_{k\to\infty} \int g_k \\ &\leq \liminf_{n\to\infty} \int f_n \end{aligned}$$

■

Theorem 15.24 (Dominated Convergence Theorem). *Given $(X, \mathscr{A}, \mu)$ a complete measure space. Let $g : X \to [0, \infty)$ a $\mathscr{A}$-measurable function with*

$$\int_X g \ d\mu < \infty$$

Let $f_n : X \to [0, \infty)$ an $\mathscr{A}$-measurable functions. Suppose $f_n \to f$ a.e. and $|f_n| \leq g$ for all n. Then f is integrable and

$$\lim \int_X f_n d\mu = \int_X f d\mu$$

Proof. We want to show that we can exchange the limit with the integral. To start, we notice $f_n \leq g$ for all $n \implies -g \leq f_n$. We can apply Fatou's lemma on the following:

$$\int_X \liminf (g - f_n) \leq \liminf \int_X g - f_n$$

$$\implies \int_X (g - \liminf f_n) \leq \int_X g - \limsup \int_X f_n$$

$$\implies \int_X \limsup f_n \geq \int_X \limsup f_n$$

Similarly, for $-g \leq f_n$, we get by Fatou's lemma:

$$\int_X \liminf (f_n + g) \leq \liminf \int_X f_n + g$$

$$\implies \int_X \liminf f_n \leq \liminf \int_X f_n$$

Therefore, since $f_n \to f$ a.e.

$$\implies \liminf \int_X f_n \leq \int_X f \leq \liminf \int_X f_n$$

Now that we can exchange the limit and the integral, we also want to demonstrate the f is L^1. But this is obvious since:

$$\int |f| d\mu = \int \lim |f_n| d\mu \leq \lim \int |f_n| d\mu \leq \lim \int g d\mu < \infty$$

Lastly, to show equality of the limits, since $f_n \to f$ in L^1, then $|f_n - f| < 2g$, so

$$\implies 0 = \int \limsup |f_n - f| \geq \limsup \int |f_n - f|$$

$$\implies \lim_{n\to\infty} |f_n - f| = 0$$

■

Note 15.25. *Updating our diagram again to include this fact, we see:*

$$\begin{tikzcd}
a.e. \arrow[r, bend left, "\mu(X)<\infty"] \arrow[d, "D.C.T."'] \arrow[dr, bend right, "\mu(X)<\infty"] & a.u \arrow[l] \arrow[d] \\
L^p \arrow[r] & \mu \arrow[ul, dashed, bend right]
\end{tikzcd}$$

Now, recall the fact that, if every subsequence of $\{a_n\} \in \mathbb{R}$ has a subsequence that converges to a, then $a_n \to a$!

Corollary 15.25.1. *If $f_n \underset{\mu}{\to} f$ and $|f_n| \leq g$ with $\int g \; d\mu < \infty$, then $f_n \to f$ in L^1.*

Proof. Define the sequence $\{a_n := \int |f_n - f|\}_n$, and let $\{a_{n_k} := \int |f_{n_k} - f|\}$ be a subsequence. Since $f_n \underset{\mu}{\to} f \implies f_{n_k} \underset{\mu}{\to} f$, then by Riesz Theorem, there exists a further subsequence $f_{n_{k_j}} \to f$ a.e.. By dominated convergence theorem, $a_{n_{j_k}} \to 0$ as $j \to \infty$. By the fact mentioned, every subsequence has a subsequence that converges to zero. Therefore, $a_n \to 0$, completing the proof. ■

Note 15.26. *Updating our diagram again to include this fact, we see:*

$$\begin{tikzcd}
a.e. \arrow[r, bend left, "\mu(X)<\infty"] \arrow[d, bend right, "D.C.T."'] \arrow[dr, bend right, "\mu(X)<\infty"] & a.u \arrow[l] \arrow[d] \\
L^p \arrow[r] & \mu \arrow[ul, dashed, bend right] \arrow[l, bend left, "D.C.T."]
\end{tikzcd}$$

Corollary 15.26.1. *Given $(X, \mathscr{A}, \mu)$ a complete measure space. Let $g : X \to [0, \infty)$ a $\mathscr{A}$-measurable function with*

$$\int_X g \; d\mu < \infty$$

Let $f_n : X \to \mathbb{R}$ a sequence of $\mathscr{A}$-measurable functions. Suppose $|f_n| \le g$. If $f_n \to f$ a.e., then $f_n \to f$ a.u.

Proof. Notice, this proof is very similar to Egorov's theorem. By Generalized Egorov, it suffices to show

$$\mu\left(\bigcup_{n=1}^{\infty}\{x \in X : |f_n(x) - f(x)| \ge \frac{1}{m}\}\right) < \infty$$

for all $m \in \mathbb{N}$. Notice, that by domination, $|f_n(x) - f(x)| \le 2g(x)$ for all x. Therefore,

$$\{x \in X : |f_n(x) - f(x)| \ge \frac{1}{m}\} \subset \{x \in X : g(x) \ge \frac{1}{2m}\}$$

$$\implies \bigcup_{n=1}^{\infty}\{x \in X : |f_n(x) - f(x)| \ge \frac{1}{m}\} \subset \{x \in X : g(x) \ge \frac{1}{2m}\}$$

$$\implies \mu\left(\bigcup_{n=1}^{\infty}\{x \in X : |f_n(x) - f(x)| \ge \frac{1}{m}\}\right) \le \mu(\{x \in X : g(x) \ge \frac{1}{2m}\})$$

Since $\int g d\mu < \infty$, then

$$\implies \mu\left(\bigcup_{n=1}^{\infty}\{x \in X : |f_n(x) - f(x)| \ge \frac{1}{m}\}\right) \le \mu(\{x \in X : g(x) \ge \frac{1}{2m}\}) < \infty$$

Therefore, we have completed the proof. ■

Note 15.27. *Updating our diagram again to include this fact, we see:*

$$\begin{tikzcd}
a.e. \arrow[r, bend left, "\mu(X)<\infty \text{ or } D.C.T."] \arrow[d, bend right, "D.C.T."'] \arrow[dr, bend right, "\mu(X)<\infty"] & a.u \arrow[l] \arrow[d, bend left] \\
L^p \arrow[r] & \mu \arrow[l, bend left, "D.C.T."] \arrow[ul, dashed, bend right, "Riesz"']
\end{tikzcd}$$

Notice that the graph is now fully connected by subsequential convergence under Dominated convergence theorem and Riesz Lemma! This means convergence in one mode guarentees subsequential convergence in another.

16 L^p Spaces

Definition 16.1. *Let $1 \le p < \infty$ and let $(X, \mathscr{A}, \mu)$ be a measure space. Then*

$$\|f\|_{L^p(X)} := \left[\int_X |f|^p \; d\mu\right]^{\frac{1}{p}}$$

Then the $\underline{L^p \text{ space}}$ is defined by:

$$L^p(X) := \{f : X \to [-\infty, \infty] : f\mathscr{A} - \textit{measurable with } \|f\|_{L^p(X)} < \infty\}$$

Further, $L^p(X)$ consists of equivalence classes. That is, given $f, g \in L^p(X)$, then

$$f \equiv g \iff f = g a.e.$$

Note 16.2. *It's easy to check that $\|\cdot\|_{L^p(X)}$ defines a semi-norm on X.*

Definition 16.3.

$$L^\infty(X) := \{f : X \to [-\infty, \infty] : f\mathscr{A} - \textit{measurable with } \|f\|_{L^\infty(X)} < \infty\}$$

where

$$\|f\|_{L^\infty(X)} := \inf\{M : \mu(\{x \in X : |f| > M\}) = 0\}$$

Note 16.4. *It's easy to check that $\|\cdot\|_{L^p(X)}$ defines a semi-norm on X.*

Lemma 16.5 (Hölder Inequality)**.** *$1 \leq p, q \leq \infty$ and let $(X, \mathscr{A}, \mu)$ be a measure space and let $f, g : X \to [-\infty, \infty]$ $\mathscr{A}$-measurable fucntions. If $f \in L^p(X), g \in L^q(X)$, where*

$$\frac{1}{p} + \frac{1}{q} = 1$$

then

$$\left|\int_X f \cdot g \; d\mu\right| \leq \int_X |fg| d\mu \leq \|f\|_{L^p} \|g\|_{L^q}$$

Question 16.6. *Where does the restriction $\frac{1}{p} + \frac{1}{q} = 1$ come from?*

We'll assume $X = \mathbb{R}$. That is, suppose we were to show,

$$\left|\int_{\mathbb{R}} f(x)g(x)dx\right| \leq \left[\int_{\mathbb{R}} |f(x)|^p dx\right]^{\frac{1}{p}} \left[\int_{\mathbb{R}} |g(x)|^q dx\right]^{\frac{1}{q}}$$

To see the above restriction, we introduce a parameter $\lambda > 0$, and then demonstrating:

$$\left|\int_{\mathbb{R}} f(\lambda x)g(\lambda x)dx\right| \leq \left[\int_{\mathbb{R}} |f(\lambda x)|^p dx\right]^{\frac{1}{p}} \left[\int_{\mathbb{R}} |g(\lambda x)|^q dx\right]^{\frac{1}{q}}$$

Define $\xi = \lambda x$,, then by change of variables, this is equivalent to

$$\int_{\mathbb{R}} f(\xi)g(\xi)\frac{d\xi}{\lambda} \leq \left[\int_{\mathbb{R}} |f(\xi)|^p \frac{d\xi}{\lambda}\right]^{\frac{1}{p}} \left[\int_{\mathbb{R}} |g(\xi)|^q \frac{d\xi}{\lambda}\right]^{\frac{1}{q}}$$

$$\implies \frac{1}{\lambda}\int_{\mathbb{R}} f(\xi)g(\xi)d\xi \leq \frac{1}{\lambda^{\frac{1}{p}+\frac{1}{q}}} \left[\int_{\mathbb{R}} |f(\xi)|^p d\xi\right]^{\frac{1}{p}} \left[\int_{\mathbb{R}} |g(\xi)|^q d\xi\right]^{\frac{1}{q}}$$

Since this integral must be true for any choice of λ, then we must impose

$$1 = \frac{1}{p} + \frac{1}{q}$$

Proof. (Of Hölder)
Case: Suppose $p = \infty, q = 1$, then it's trivial to see that:

$$\int_X |fg|d\mu \leq \|f\|_\infty \int_X |g|d\mu = \|f\|_{L^\infty(X)} \|g\|_{L^1(X)}$$

Case: Suppose $1 < p, q < \infty$. Therefore,

$$\|f\|_{L^p(X)} < \infty \qquad \|g\|_{L^q(X)} < \infty$$

Now, observe the following a equivalent:

$$\int |fg|d\mu \leq \|f\|_{L^p(X)} \|g\|_{L^q(X)} \iff \int \underbrace{\frac{|f|}{\|f\|_{L^p(X)}}}_{F} \underbrace{\frac{|g|}{\|g\|_{L^q(X)}}}_{G} d\mu \leq 1$$

Then it follows F, G are normalized:

$$\|F\|_{L^p(X)} = 1 \qquad \|G\|_{L^q(X)} = 1$$

So now we need to prove given $F \in L^p, G \in L^q$ such that

$$\|F\|_{L^p(X)} = 1 \qquad \|G\|_{L^q(X)} = 1$$

then we need to show

$$\int FG \leq 1$$

We can do this by rewriting,

$$\begin{aligned}
FG &= e^{\ln FG} \\
&= e^{\frac{1}{p}\ln F^p + \frac{1}{q}\ln F^q} \\
&\frac{1}{p}e^{\ln F^p} + \frac{1}{q}e^{\ln G^q} && \text{By convexity} \\
&= \frac{1}{p}F^p + \frac{1}{q}G^q && \text{This result is known as Young's Inequalty}
\end{aligned}$$

Therefore,

$$\int FG \leq \int \left(\frac{1}{p}F^p + \frac{1}{q}G^q\right) = \frac{1}{p}\|F\|^p_{L^p(X)} + \frac{1}{q}\|G\|^q_{L^q(X)} = \left(\frac{1}{p} + \frac{1}{q}\right) \cdot 1 = 1$$

■

Corollary 16.6.1 (Triangle Inequality in $L^p(X)$)**.**

$$\|f + g\|_{L^p(X)} \leq \|f\|_{L^p(X)} + \|g\|_{L^q(X)}$$

Proof. *Case:* When $p = \infty$, the conclusion is immediate.
Case: When $1 \leq p < \infty$, then

$$\begin{aligned}
\|f + g\|^p_{L^p} &= \int |f + g|^p d\mu \\
&= \int |f + g|^{p-1}|f + g|d\mu \\
&\leq \int \underbrace{|f + g|^{p-1}}_{L^q} \underbrace{|f|}_{L^p} d\mu + \int \underbrace{|f + g|^{p-1}}_{L^q} \underbrace{|g|}_{L^p} d\mu \\
&= \left(\int |f + g|^{(p-1)q}\right)^{\frac{1}{q}} \|f\|_{L^p} + \left(\int |f + g|^{(p-1)q}\right)^{\frac{1}{q}} \|g\|_{L^p} && \text{Hölder} \\
&= \|f + g\|^{\frac{p}{q}}_{L^p} (\|f\|_{L^p} + \|g\|_{L^p})
\end{aligned}$$

Dividing by $\|f+g\|_{L^p}^{\frac{p}{q}}$ on both sides gives us

$$\frac{1}{p}+\frac{1}{q}=1 \implies 1+\frac{p}{q}=p \implies p-\frac{p}{q}=1$$
$$\implies \|f+g\|_{L^p} \le \|f\|_{L^p}+\|g\|_{L^p}$$

■

Theorem 16.7. *$L^p(X)$ is a Banach Space.*

Proof. We simply need to show that the space is complete. Recall the fact that a normed linear space is complete if and only if absolute convergence implies convergence in L^p. To start, suppose

$$\sum_i |f_i| \underset{L^p}{\longrightarrow} \tilde{g} \in L^p$$

Define the function $g := (\sum_i |f_i|)^p \in L^1$. Notice, $g \ge 0$. By these two facts, then we know g is finite almost everywhere. On the other hand, we can define our convergence candidate f by

$$f(x)=\begin{cases} \sum_i f_i(x) & g(x)<\infty \\ 0 & \text{Otherwise} \end{cases}$$

It's obvious to see the partial sums

$$\left\{ S_N(x) := \sum_{i=1}^{N} f_i(x) \right\}_N \to f(x) a.e.$$

as well as $|S_N - f|^p \le 2g \in L^1$. Lastly, $|S_N - f|^p \to 0$ a.e. By Dominated convergence theorem, then

$$\lim_{N\to\infty} \|S_N - f\|_{L^p} = \left\| \lim_{N\to\infty} S_N - f \right\|_{L^p} = 0$$

■

Remark 16.8. *In general, $f \in L^p(X) \not\Longrightarrow f \in L^q(X)$ where $p \ne q$*

Example 16.9. *$\frac{1}{x^\alpha}$ depends sensitively on the exponent p, q.*

Remark 16.10. $f \in \bigcup_{1\le p<\infty} L^p(X) \not\Longrightarrow f \in L^\infty(X)$

Example 16.11. *$f : [0,1] \to \mathbb{R}$ defined by $f(x) = \ln(\frac{1}{x})$*

Theorem 16.12. *Suppose $\mu(X) < \infty$, if $1 \le p < q \le \infty$, then $L^p(X) \subset L^q(X)$.*

Proof. If $f \in L^\infty$, then

$$\left[\int_X |f|^p\right]^{\frac{1}{p}} \le \|f\|_{L^\infty} \mu(X)^{\frac{1}{p}} < \infty \implies L^\infty(X) \subset L^p(X)$$

If $f \in L^q(X)$ with $q > p$, then

$$\frac{q}{p} > 1$$

Therefore, we can use the Holder Complement $(\frac{q}{p})^*$ such that

$$\frac{1}{\frac{q}{p}} + \frac{1}{(\frac{q}{p})^*} = 1$$

To manipulate the following integral using Hölder's Inequality:

$$\left|\int_X \underbrace{|f|^p}_{L^{\frac{q}{p}}} \mathcal{X}_X\right| \leq \||f|^p\|_{L^{\frac{q}{p}}} \mu(X)^{\frac{1}{(q/p)^*}} \leq \|f\|^p_{L^{\frac{q}{p}}} \mu(X)^{\frac{1}{(q/p)^*}}$$

Therefore, $f \in L^q(X)$ ■

Theorem 16.13. *When* $\mu(X) < \infty$,

$$L^\infty(X) = \bigcap_{1 \leq p < \infty} L^p(X)$$

and

$$\|f\|_{L^\infty} = \limsup_{p \to \infty} \|f\|_{L^p}$$

Proof. Observe,

$$\limsup_{p \to \infty} \left[\int |f|^p d\mu\right]^{\frac{1}{p}} \leq \lim_{p \to \infty} \|f\|_{L^\infty} \mu(X)^{\frac{1}{p}} = \|f\|_{L^\infty}$$

To show the other side of the inequality, pick $0 \leq c < \|f\|_{L^\infty}$ so

$$\mu(\{x \in X : |f(x)| > c\}) \neq 0$$

We want to show

$$\limsup_{p \to \infty} \|f\|_{L^p} \geq \|f\|_{L^\infty}$$

It suffices to show that

$$\limsup_{p \to \infty} \|f\|_{L^p} \geq c$$

We can do so by:

$$\left(\int_X |f|^p \, d\mu\right)^{\frac{1}{p}} \geq \left(\int_{\{x \in X : |f(x)| > c\}} |f|^p \, d\mu\right)^{\frac{1}{p}} \geq c\mu(\{x \in X : |f(x)| > c\})^{\frac{1}{p}} > 0$$

Therefore, as $p \to \infty$, then

$$\limsup_{p \to \infty} \|f\|_{L^p} \geq c$$

Lastly, we can take any sequence $\{c_i\} \to \|f\|_{L^\infty}$ which will prove the second inequality. ■

16.1 Interpolation Theorems

Theorem 16.14. *Suppose we are given* $(X, \mathscr{A}, \mu)$ *a measures space and let* $1 \leq p < q \leq \infty$. *If* $f \in L^p(X) \cap L^q(X)$, *then* $f \in L^r$ *where*

$$\frac{1}{r} = \frac{1-\lambda}{p} + \frac{\lambda}{q}$$

with $\lambda \in [0, 1]$. *We can also control the norm in* L^r *by*

$$\|f\|_{L^r} \leq \|f\|_{L^p}^{1-\lambda} \|f\|_{L^q}^{\lambda}$$

Proof. If we want to control the L^r norm in terms of the L^p, L^q norms. Since

$$r = (1-\lambda)r + \lambda r$$

Then we we can transform the integral:

$$\begin{aligned}
\int |f|^r &= \int |f|^{(1-\lambda)r+\lambda r} \\
&= \int |f|^{(1-\lambda)r}|f|^{\lambda r} \\
&\leq \left\| |f|^{(1-\lambda)r} \right\|_{L^{p_1}} \left\| |f|^{\lambda r} \right\|_{L^{q_1}} && \text{Hölder with Exponents } \underbrace{\frac{1}{p[(1-\lambda)r]^{-1}}}_{p_1} + \underbrace{\frac{1}{q[\lambda r]^{-1}}}_{q_1} = 1 \\
&= \left[\int |f|^{(1-\lambda)rp_1}\right]^{\frac{1}{p_1}} \left[\int |f|^{\lambda r q_1}\right]^{\frac{1}{q_1}} \\
&= \left[\int |f|^p\right]^{\frac{(1-\lambda)r}{p}} \left[\int |f|^q\right]^{\frac{\lambda r}{q}}
\end{aligned}$$

$$\begin{aligned}
&\implies \left[\int |f|^r\right]^{\frac{1}{r}} \leq \left[\int |f|^p\right]^{\frac{1-\lambda}{p}} \left[\int |f|^q\right]^{\frac{\lambda}{q}} \\
&\iff \|f\|_{L^r} \leq \|f\|_{L^p}^{1-\lambda} \|f\|_{L^q}^{\lambda}
\end{aligned}$$

■

Notice that we can control the norm in L^r by the norms of L^p and L^q. We want to generalize this notion to operator norms as well through the Riesz-Thorin Interpolation Theorem.

Theorem 16.15 (Riesz-Thorin Interpolation Theorem)**.** *Suppose we are given two $(X, \mathscr{A}_1, \mu_1), (Y, \mathscr{A}_2, \mu_2)$ σ-finite measure spaces. Also, suppose*

$$1 \leq p_1 \leq p_2 \leq \infty$$

$$1 \leq q_1 \leq q_2 \leq \infty$$

Lastly, let

$$T : L^{p_1}(X) + L^{p_2}(X) \to L^{q_1}(Y) + L^{q_2}(Y)$$

be a linear operator such that the restrictions $T|_{L^{p_1}(X)} : L^{p_1}(X) \to L^{q_1}(Y)$, $T|_{L^{p_2}(X)} : L^{p_2}(X) \to L^{q_2}(Y)$ are well-defined and bounded. Then T is a bounded linear operator

$$T : L^{p_\lambda}(X) \to L^{q_\lambda}(Y)$$

where $\lambda \in [0,1]$ and

$$\frac{1}{p_\lambda} = \frac{1-\lambda}{p_1} + \frac{\lambda}{p_2}$$

$$\frac{1}{q_\lambda} = \frac{1-\lambda}{q_1} + \frac{\lambda}{q_2}$$

Moreover, the norm can be controlled as:

$$\|T\|_{L^{p_\lambda} \to L^{q_\lambda}(Y)} \leq \|T\|_{L^{p_1} \to L^{q_1}}^{1-\lambda} \|T\|_{L^{p_2} \to L^{q_2}}^{\lambda}$$

We shall prove this later.

Claim: Recall, $L^p \cap L^q \subset L^r$ for some $p < r < q$. But suppose $f \in L^r$. Then by the previous theorem, we can decompose f by

$$f = f_1 + f_2$$

with $f_1 \in L^p$ and $f_2 \in L^q$. Moreover,

$$L^r \subset L^p + L^q$$

Proof. Let $1 \le p < r < q \le \infty$. Then

$$f = \underbrace{f\mathcal{X}_{\{|f|\le 1\}}}_{f_2 \in L^q} + \underbrace{f\mathcal{X}_{\{|f|>1\}}}_{f_1 \in L^p}$$

Observe, we want to pair small quantities with big exponent and vice-versa:

$$\int |f\mathcal{X}_{\{|f|\le 1\}}|^q = \int_{\{|f|\le 1\}} |f|^q \le \int_{\{|f|\le 1\}} |f|^r \le \|f\|_{L^r}^r \implies f_2 \in L^q$$

$$\int |f\mathcal{X}_{\{|f|>1\}}|^p = \int_{\{|f|>1\}} |f|^p \le \int_{\{|f|>1\}} |f|^r \le \|f\|_{L^r}^r \implies f_1 \in L^p$$

■

16.2 Weak L^p Spaces

Definition 16.16. *Given $(X, \mathscr{A}, \mu)$ a measure space with $1 \le p < \infty$. We define the space*

$$L^{p,weak}(X)$$

as the set of functions $f \in L^{p,weak}(X)$ provided $f : X \to \mathbb{R}$ is a $\mathscr{A}$-measurable function and

$$\mu(\{x \in X : |f(x)| > \lambda\}) \le \frac{C^p}{\lambda^p}$$

for all $\lambda > 0$ and some choice constant C_λ.

Lemma 16.17. $L^p(X) \subset L^{p,weak}(X)$

Proof. If $f \in L^p(X)$ then as a consequence of the Chebyshev inequality,

$$\mu(\{x \in X : |f(x)| > \lambda\}) \le \frac{\|f\|_{L^p}^p}{\lambda^p}$$

■

Definition 16.18. *The distribution function $m_\mu(\lambda) : \mathbb{R}^+ \to [0, \infty]$ by*

$$m_\mu(\lambda) := \mu(\{x \in X : |f(x)| > \lambda\})$$

Further, we can define a norm for this space via abuse of notation (as Kevin says):

Definition 16.19.

$$\|f\|_{L^{p,weak}(X)} = \sup_{\lambda > 0} \left\{ \lambda m_\mu(\lambda)^{\frac{1}{p}} \right\}$$

Note 16.20. *This is NOT a norm. It's a quasi-norm because this fails to satisfy the triangle inequality.*

Example 16.21. *Take $X = (0,1)$ and $f = \frac{1}{x}$ and $g(x) = f(1-x)$. Then*

$$\|f+g\|_{L^{p,weak}(X)} > \|f\|_{L^{p,weak}(X)} + \|g\|_{L^{p,weak}(X)}$$

We'll still call this a norm even though it's not because it's fun to give people aneurysms.
Now let's compute the L^p norm using the distribution function. First, we'll recall a few facts from measure theory:

Lemma 16.22. $1 \le p < \infty, \|f\|_{L^p}^p = \int_X |f|^p \; d\mu = p\int_0^\infty \lambda^{p-1} m_\mu(\lambda) d\lambda$

Proof.

$$\begin{aligned} p\int_0^\infty \lambda^{p-1} m_\mu(\lambda) d\lambda &= p\int_0^\infty \lambda^{p-1}\mu(\{x\in X : |f(x)| > \lambda\}) d\lambda \\ &= p\int_0^\infty \lambda^{p-1}\int_{|f|>\lambda} d\mu d\lambda \\ &= \int_X\int_0^{|f|} p\lambda^{p-1} d\lambda d\mu && \text{Fubini} \\ &= \int_X |f|^p d\mu \\ &= \|f\|_{L^p}^p \end{aligned}$$

■

Theorem 16.23 (Marcin Liewicz Interpolation, 1939)**.** *Given $1 \le p < q < \infty$, suppose T is subadditive operator*

$$T : L^p(X) + L^q(X) \to L^{p,weak}(X) + L^{q,weak}(X)$$

If the restrictions $T|_{L^p(X)} : L^p(X) \to L^{p,weak}(X)$, $T|_{L^q(X)} : L^q(X) \to L^{q,weak}(X)$ are well-defined and bounded, then $T : L^r \to L^r$ for $p < r < q$.

Observe, this means that the quantity should be controlled by:

$$\mu(\{x \in X : |Tf| > \lambda\}) \le \frac{\|f\|_{L^p}^p}{\lambda^p}$$

$$\mu(\{x \in X : |Tf| > \lambda\}) \le \frac{\|f\|_{L^q}^q}{\lambda^q}$$

which means we can control the norm:

$$\|Tf\|_{L^r} \le C\,\|f\|_{L^r}$$

Proof. We want to show $\|Tf\|_{L^r}^r \le C\,\|f\|_{L^r}^r$ for some constant C. Observe,

$$\int |Tf|^r = r\int_0^\infty \lambda^{r-1}\mu\{x \in X : |Tf(x)| > \lambda\}\; d\lambda$$

Now, certainly if we split

$$f = \underbrace{f\mathcal{X}_{\{|f|\le 1\}}}_{f_1} + \underbrace{f\mathcal{X}_{\{|f|>1\}}}_{f_2}$$

Then by the sub-additivity of T, we see:

$$\lambda < |Tf| = |T(f_1+f_2)| \le |Tf_1| + |Tf_2|$$

Therefore,

$$\mu\{x \in X : |Tf(x)| > \lambda\} \leq \mu\{x \in X : |Tf_1(x)| > \frac{\lambda}{2}\} + \mu\{x \in X : |Tf_2(x)| > \frac{\lambda}{2}\} \leq \frac{\|f_1\|_{L^q}^q}{\lambda^p} + \frac{\|f_2\|_{L^p}^p}{\lambda^p}$$

Therefore, we can use these estimates to bound our integral:

$$\begin{aligned}
\int |Tf|^r &= r\int_0^\infty \lambda^{r-1}\mu\{x \in X : |Tf(x)| > \lambda\}\, d\lambda \\
&\leq r\int_0^\infty \lambda^{r-1}\left(\frac{\|f_1\|_{L^q}^q}{\lambda^p} + \frac{\|f_2\|_{L^p}^p}{\lambda^p}\right)\, d\lambda \\
&\leq \underbrace{r\int_0^\infty \lambda^{r-1}\frac{\|f_1\|_{L^q}^q}{\lambda^p}\, d\lambda}_{I_1} + \underbrace{r\int_0^\infty \lambda^{r-1}\frac{\|f_2\|_{L^p}^p}{\lambda^p}\, d\lambda}_{I_2}
\end{aligned}$$

Evaluating each integral separately,

$$\begin{aligned}
I_1 &= r\int_0^\infty \lambda^{r-1}\frac{\|f_1\|_{L^q}^q}{\lambda^p}\, d\lambda \\
&= r\int_0^\infty \lambda^{r-q-1}\int_{|f|\leq\lambda} |f|^q d\mu\; d\lambda \\
&= \int |f|^q \int_{|f|}^\infty r\lambda^{r-q-1} d\lambda d\mu && \text{Fubini} \\
&= \int |f|^q \frac{r}{q-r}|f|^{r-q} d\mu && r-q-1 < -1 \implies \text{Integrable!} \\
&= \frac{r}{q-r}\int |f|^r d\mu
\end{aligned}$$

$$\begin{aligned}
I_2 &= r\int_0^\infty \lambda^{r-1}\frac{\|f_2\|_{L^p}^p}{\lambda^p}\, d\lambda \\
&= \int |f|^p \int_0^{|f|} r\lambda^{r-1-p} d\lambda d\mu && \text{Fubini} \\
&= \int |f|^p \frac{r}{r-p}|f|^{r-p} \\
&= \frac{r}{r-p}\int |f|^r d\mu
\end{aligned}$$

Therefore,

$$\|Tf\|_{L^r}^r \leq \left(\frac{r}{q-r} + \frac{r}{r-p}\right)\|f\|_{L^r}^r$$

■

16.3 Approximating L^p Functions

Definition 16.24. *A function of the form*

$$\phi(x) = \sum_{i=1}^N a_i \mathcal{X}_{A_i}(x)$$

with $A_i \subset X$ measurable with $\mu(A_i) < \infty, |a_i| < \infty$ is a simple function.

Lemma 16.25. *Simple functions are dense in L^p.*

That is, given an L^p function, we can generate a sequence of simple functions to approximate the function.

Proof. Let $f \in L^p(X)$. Without loss of generality, we may assume $f \geq 0$. (Otherwise we can write $f = f^+ - f^-$) We need to construct a sequence of simple functions such that

$$s_1 \leq s_2 \leq \ldots \leq f$$

monotonically increasing. Define the sets:

$$E_{k,n} := \{x \in X : \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n}\}$$

from $1 \leq k \leq (n+1)2^n$. Taking the union we see:

$$\bigcup_{n \geq 1} \bigcup_{1 \leq k \leq (n+1)2^n} E_{k,n} = X$$

Now define the following simple function:

$$s_n := \sum_{k=1}^{(n+1)2^n} \frac{k-1}{2^n} \mathcal{X}_{E_{k,n}}$$

Clearly the following are true of our construction:

- $s_n \leq s_{n+1}$.
- $s_n \to f$ a.e.
- $s_n \leq f$

Since $f \in L^p \implies s_n \in L^p$. Therefore, we see that since we can dominate the quantity:

$$|s_n - f|^p \leq 2^p f^p \in L^1$$

then by dominated convergence theorem, we know:

$$|s_n - f|^p \to 0 \ a.e. \implies \lim_{n \to \infty} \|s_n - f\|_{L^p} = \left\| \lim_{n \to \infty} s_n - f \right\|_{L^p} = 0$$

■

Note 16.26. *We used the fact that our construction of s_n yields the approximation:*

$$\|f - s_n\|_{L^\infty} \leq \frac{1}{2^n}$$

and therefore, $s_n \to f$ in L^∞.

Definition 16.27. *On the Lebesgue measure space $(X, \mathscr{B}(X), \lambda)$, we define the step function of the form*

$$\phi(x) = \sum_{i=1}^{N} a_i \mathcal{X}_{Q_i}(x)$$

where Q_i are pairwise disjoint cubes of the form:

$$Q_i := \prod_{j=1}^{n} [\alpha_j, \alpha_j + \ell)$$

with $\alpha_i, \ell \in \mathbb{R}^+$ and $|a_i| < \infty$.

Question 16.28. *We know that*

$$\{ \text{ step functions } \} \subset \{ \text{ simple functions } \}$$

Can we maintain density in L^p spaces?

Answer: Almost.

Remark 16.29. *Steps functions are NOT dense in $L^\infty(\mathbb{R}^n)$.*

Example 16.30. *On $\mathbb{R}$, there exists a Lebesgue measurable set $A \subset \mathbb{R}$ such that*

$$\begin{cases} \lambda(A \cap I) > 0 & \text{for all nonempty open intervals } I \\ \lambda(A^c \cap I) > 0 & \text{otherwise} \end{cases}$$

Consider the function $\mathcal{X}_A \in L^\infty$.
Claim: *X_A cannot be approximate uniformly by step functions.*
Let $\phi = \sum_{i=1}^\infty c_i X_{I_i}$. Then for any choice of i

$$\|\phi - \mathcal{X}_A\|_{L^\infty} \geq \|c_i \mathcal{X}_{I_i} - \mathcal{X}_{A_i}\|_{L^\infty} \geq \max\{|1 - c_i|, |c_i|\} \geq \frac{1}{2}$$

Lemma 16.31. *Step functions are dense in $L^p(\mathbb{R}^n, \lambda)$ for $1 \leq p < \infty$.*

Proof. Suppose $f \in L^p(\mathbb{R}^n) \implies f \sim \sum_{i=1}^N a_i \mathcal{X}_{A_i}$ a simple function. So it is enough to show the $X_A \sim \sum_{i=1}^N a_i X_{Q_i}$ a step function for any measurable set A. To do so, we will show equivalence of our hypercubes in $\mathbb{R}^n$ with open sets under step functions.
($\Rightarrow$) If the Q_i are allowed to be any open set, then the conclusion is immediate. Why? Because if we are given any measurable set A, then there exists an open set $\mathcal{O} \in \mathbb{R}^n$ such that

$$\mathcal{O} \supset A \text{ such that } \lambda(\mathcal{O} \setminus A) \leq \epsilon^p$$

$$\implies \|\mathcal{X}_\mathcal{O} - \mathcal{X}_A\|_{L^p} < \epsilon$$

($\Leftarrow$) Let $\mathcal{O} \subset \mathbb{R}^n$ be an arbitrary open set. We will show

$$\mathcal{X}_\mathcal{O} \sim \sum_{i=1}^N \mathcal{X}_{Q_i}$$

where Q_i are pairwise disjoint with finite measure such that

$$\left\| \mathcal{X}_\mathcal{O} - \sum_{i=1}^N \mathcal{X}_{Q_i} \right\|_{L^p(\mathbb{R}^n)}$$

as $N \to \infty$. In other words, we will show that the open set can be covered in pairwise disjoint cubes:

$$\mathcal{O} = \bigcup_{i=1}^\infty Q_i$$

with Q_i pairwise disjoint.

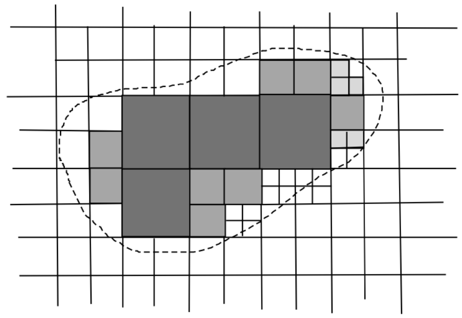


Figure 1.5. Each of the intervals contained in G are retained; the rest are subdivided.

Consider the following algorithm for approximating an open set with cubes:

1. Start with the grid of unit cubes in $\mathbb{R}^n$ and let

$$E_0 = \{ \text{ unit cubes } Q \text{ such that } Q \cap \mathcal{O} \neq \emptyset\}$$

2. A cube Q is <u>OK</u> if $Q \subset \mathcal{O}$
3. Let $E_0' := \{Q \in E_0 : Q \text{ is OK } \}$
4. If $Q \in E_0 \setminus E_0'$, then we bisect Q into 2^n congruent subcubes.
5. Let $E_{n+1} := \{ \text{ subcubes } \in E_n \setminus E_n'\}$.
6. Repeat from step 3 until $E_n \setminus E_n' = \emptyset$.
7. Lastly, we conclude that

$$\mathcal{O} = \bigcup_{i=0}^{n} \bigcup_{Q \in E_i'} Q$$

This equivalence proves step functions are dense in L^p. ∎

It should come as no surprise that continuous functions carry this same property:

Corollary 16.31.1. *Continuous functions are dense in $L^p(\mathbb{R}^n)$ with $1 \leq p < \infty$*

Definition 16.32. *The <u>support</u> of f is defined as the set*

$$supp(f) := \{x \in X : f(x) \neq 0\}$$

Definition 16.33. *$C_c(\mathbb{R}^n) =$ the space of continuous functions of compact support.*

Lemma 16.34. *The set $C_c(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$*

Proof. Let $f \in L^p(\mathbb{R}^n)$, we define the function

$$f^{[n]}(x) := \begin{cases} f(x) & |x| \leq n, |f| \leq n \\ 0 & \text{otherwise} \end{cases}$$

Clearly,

$$\left\| f^{[n]} - f \right\|_{L^p(\mathbb{R}^n)} \to 0$$

as $n \to \infty$. Therefore, all we need to do is approximate $f^{[n]}$ by continuous functions.
Case: Suppose f already has compact support. Then $f \in L^p \implies f$ can be approximated by step functions

$$\left\| f - \sum_{i=1}^{N} a_i \mathcal{X}_{Q_i} \right\|_{L^p} < \epsilon$$

Let C_i be a closed cube such that

- $C_i \subset Q_i$
- $\mu(Q_i \setminus C_i \leq \frac{\frac{\epsilon^p}{2^i}}{2^p \max\{1,|a_i|^p\}}$

To do this, we will use the Urysohn function for each pair (Q_i, C_i) :

$$g_i(x) = \frac{dist(x, Q_i^c)}{dist(x, Q_i^c) + dist(x, C_i)}$$

a continuous function. Therefore,

$$\begin{aligned}
\left\| \sum_{i=1}^{N} a_i \mathcal{X}_{Q_i} - \sum_{i=1}^{N} a_i g_i \right\|_{L^p} &\leq \sum_{i=1}^{N} |a_i|^p \int_{Q_i} |\mathcal{X}_{Q_i} - q_i|^p d\lambda \\
&= \sum_{i=1}^{N} |a_i|^p 2^p \lambda(Q_i \setminus C_i) \\
&\leq \sum_{i=1}^{N} |a_i|^p 2^p \frac{\frac{\epsilon^p}{2^i}}{2^p \max\{1, |a_i|^p\}} \\
&\leq \epsilon^p \sum_{i=1}^{N} \frac{1}{2^i} \leq \epsilon^p
\end{aligned}$$

■

16.4 Dual Characterization of $L^p(X)$

Theorem 16.35. *Given a $(X, \mathscr{A}, \mu)$ a measure space and $1 \leq p < \infty$ with*

$$\frac{1}{p} + \frac{1}{q} = 1$$

Then

$$\|f\|_{L^p(X)} = \sup_{\|g\|_{L^q(X)}=1} \int_X f \cdot g d\mu$$

If $p = \infty$ and X is σ-finite, then

$$\|f\|_{L^\infty(X)} = \sup_{\|g\|_{L^1(X)}=1} \int_X f \cdot g d\mu$$

Proof. *Case:* Suppose $p < \infty$. It's obvious that provided $\|g\|_{L^q} = 1$, then

$$|\int_X f g d\mu| \leq \|f\|_{L^p} \|g\|_{L^q} = \|f\|_{L^p}$$

by Hölder's inequality. Therefore, the supremem of both side maintains this:

$$\sup_{\|g\|_{L^q}=1} |\int_X f g d\mu| \leq \sup_{\|g\|_{L^q}=1} \|f\|_{L^p} \|g\|_{L^q} = \sup_{\|g\|_{L^q}=1} \|f\|_{L^p}$$

Therefore, we would like to check that other inequality. To do so, we need tp pick g such that $g \in L^q$ with $\|g\|_{L^q} = 1$ and

$$\int_X fg \; d\mu = \|f\|_{L^p}$$

- *Subcase:* $q = \infty$. Then such a g must satisfy

  $$\int_X fg d\mu = \|f\|_{L^1} = \int_X |f| d\mu$$

  Therefore, choosing

  $$g = sign(f)$$

  implies

  $$\sup_{\|g\|_{L^q}=1} |\int_X fg d\mu| \geq \|f\|_{L^1}$$

- *Subcase:* $q < \infty$. We need

  $$\int_X fg d\mu = \|f\|_{L^p} = \left(\int_X |f|^p d\mu\right)^{1/p}$$

  Notice, must lie within the space $g \in L^q = L^{\frac{p}{p-1}}$. Therefore, choosing

  $$\tilde{g} = sign(f) \cdot |f|^{p-1} \in L^q$$

  satisfies the integration property. However, this needs to be normalized: therefore the choice

  $$g = \frac{sign(f) \cdot |f|^{p-1}}{\|f^{p-1}\|_{L^q}} \in L^q$$

  implies

  $$\begin{aligned}
  \int_X fg d\mu &= \int_X f \cdot \frac{sign(f) \cdot |f|^{p-1}}{\|f^{p-1}\|_{L^q}} d\mu \\
  &= \frac{1}{\|f^{p-1}\|_{L^q}} \int_X |f||f|^{p-1} d\mu \\
  &= \frac{1}{\|f^{p-1}\|_{L^q}} \|f\|_{L^p}^p \\
  &= \frac{1}{\|f\|_{L^p}^{p-1}} \|f\|_{L^p}^p \\
  &= \|f\|_{L^p}
  \end{aligned}$$

Case: Suppose $p < \infty$. We now need to assume the σ-finitedness of the space in order to achieve equality. That is,

$$X = \bigcup_{n=1}^{\infty} X_n \quad \text{with} \quad \mu(X_n) < \infty$$

By Hölder's Inequality again, for any $f \in L^\infty$

$$\|f\|_{L^\infty} \geq \sup_{\|g\|_{L^1}=1} \int_X f \cdot g \; d\mu$$

Assume $\|f\|_{L^\infty} = 0$ (otherwise the conclusion is immediate). Fix $\epsilon > 0$ and define

$$A_\epsilon := \{x \in X : |f(x)| \geq \|f\|_{L^\infty} - \epsilon\}$$

So to show the reverse inequality, we can define $g \in L^1(X)$ by

$$g = sign(f)\frac{\mathcal{X}_{A_\epsilon \cap X_n}}{\mu(A_\epsilon \cap X_n)}$$

in which $\mu(A_\epsilon \cap X_n) \neq 0$ for all n. Clearly this function g is within $L^1(X)$ with $\|g\|_{L^1} = 1$ and

$$\begin{aligned}\int_X f \cdot g d\mu &= \frac{1}{\mu(A_\epsilon \cap X_n)} \int_{A_\epsilon \cap X_n} sign(f) \cdot f d\mu \\ &= \frac{1}{\mu(A_\epsilon \cap X_n)} \int_{A_\epsilon \cap X_n} |f| d\mu \\ &\geq \|f\|_{L^\infty} - \epsilon\end{aligned}$$

Since this is true for all L^1 functions, then it must be true that:

$$\|f\|_{L^\infty} \leq \sup_{\|g\|_{L^1}=1} \int_X f \cdot g \ d\mu$$

■

To summarize, if for $1 \leq p < \infty$ then

$$f \in L^p(X) \iff \sup_{\|g\|_{L^q}=1} \int f \cdot g d\mu < \infty$$

And if X is σ-finite, then

$$f \in L^\infty(X) \iff \sup_{\|g\|_{L^1}=1} \int f \cdot g d\mu < \infty$$

16.5 Riesz Representation Theorem

Definition 16.36. *A bounded linear functional on $L^p(X)$ is a mapping $\ell : L^p(X) \to \mathbb{R}$ with*

$$\ell(af + bg) = a\ell(f) + b\ell(g)$$

and

$$|\ell(f)| \leq C \|f\|_{L^p}$$

for some constant C. Moreover, we define

$$\|\ell\| := \inf\{C : |\ell(f)| \leq C \|f\|_{L^p} \, \forall f \in L^p\}$$

Example 16.37. *Given $g \in L^q(X)$, we can define the linear functional*

$$\ell_g : L^p \to \mathbb{R}$$

$$\ell_g(f) := \int_X f \cdot g \ d\mu$$

Notice $\|\ell_g\| = \|g\|_{L^q}$

Question 16.38. *What are all the bounded linear functionals on $L^p(X)$?*

Definition 16.39. *The set of all bounded linear functionals on L^p is called the dual of $L^p(X)$, denoted $(L^p(X))^*$.*

Theorem 16.40 (Riesz Representation Theorem on $L^p(X)$)**.** *If $1 \leq p < \infty$ and X is σ-finite, then*

$$(L^p(X))^* = L^q(X)$$

Moreover, suppose $\ell \in (L^p(X))^*$. Then there exists a unique $g \in L^q(X)$ such that

$$\ell(f) = \ell_g(f) = \int_X f \cdot g d\mu$$

Before we prove this theorem, we need another result to help us.

Definition 16.41. *Suppose we have a measurable space $(X, \mathscr{A})$ equipped with two finite measures μ, ν. ν is absolutely continuous with respect to μ, denoted $\nu \ll \mu$, provided*

$$\mu(B) = 0 \implies \nu(B) = 0$$

Theorem 16.42 (Radon-Nikodym)**.** *Suppose we have a measurable space $(X, \mathscr{A})$ equipped with two finite measures μ, ν. Further, suppose*

$$\nu \ll \mu$$

Then there exists a measurable, nonnegative, integrable function f such that

$$\nu(A) := \int_A f \ d\mu$$

for any $A \in \mathscr{A}$.

Proof. Recall that on a Hilbert Space H, for any $\ell \in H^*$, then Riesz Representation on Hilbert spaces given us the result that there must exist a unique $g \in H$ such that

$$\ell(x) = \langle x, g \rangle$$

Therefore, we define $\pi := \mu + \nu$ and define the Hilbert space $L^2(X, \mathscr{A}, \mu)$ with inner product

$$\langle f, g \rangle := \int_X f \cdot g \ d\pi$$

Now define the functional

$$\ell(f) := \int_X f d\nu$$

Claim: $\ell \in (L^2(X, \mathscr{A}, \mu))^*$

$$\begin{aligned}
|\ell(f)| &= |\int_X f d\nu| \\
&\leq \int_X |f| d\nu \\
&\leq \int_X |f| d\pi \\
&\leq \|f\|_{L^2} \cdot \sqrt{\pi(X)} && \text{Hölder}
\end{aligned}$$

Since the measure space is finite, we see:

$$\implies \|\ell\| < \sqrt{\pi(X)} < \infty$$

Therefore, we have shown that $\ell \in (L^2(X, \mathscr{A}, \mu))^*$. Therefore, applying the familiar Riesz Representation Theorem, we know there exists a unique $g \in L^2(X, \mathscr{A}, \mu)$ such that

$$\int_X f d\nu = \int f \cdot g d\pi = \int_X fg(d\mu + d\nu)$$

$$\iff \int_X f(1-g)d\nu = \int_X fgd\mu$$

Now we want $f \cong \frac{1}{1-g}$ in order to integrate to isolate ν. We have to be careful that $g \neq 1$. Therefore, we define three sets and analyze the problem on each of these sets:

$$B = \{x \in X : g(x) \geq 1\}$$

$$N = \{x \in X : g(x) \leq 0\}$$

$$G = \{x \in X : 0 < g(x) < 1\}$$

- Letting $f = \mathcal{X}_B$ Then

$$\int_X fgd\mu = \int_X f(1-g)d\nu$$

$$\iff \int_B \underbrace{g}_{\geq 1} d\mu = \int_B \underbrace{(1-g)}_{\leq 0} d\nu$$

  Therefore, in order to maintain equality, it must follow that:

$$\iff \int_B gd\mu = 0 = \int_B (1-g)d\nu$$

  Therefore,

$$0 = \int_B gd\mu \geq \int_B d\mu = \mu(B) \implies \mu(B) = 0$$

  By the absolute continuity of $\nu \ll \mu \implies \nu(B) = 0$. Therefore, we reach our first conclusion that

$$\mu(B) = \nu(B) = 0$$

- Let $f = \mathcal{X}_N$. By the same logic

$$\int_N \underbrace{(1-g)}_{\geq 1} d\nu = \int_N \underbrace{g}_{<0} d\mu$$

  Therefore, by the same logic above, we see

$$\implies \mu(N) = \nu(N) = 0$$

Therefore, we see that since $X = B \cup N \cup G$, then

$$\int_X f(1-g)d\nu = \int_G f(1-g)d\nu = \int f \cdot gd\mu$$

Therefore, for any $A \in \mathscr{A}$,

$$\begin{aligned}\nu(A) &= \int_X \mathcal{X}_A d\nu \\ &= \int_X \frac{\mathcal{X}_A}{1-g}(1-g)d\nu \\ &= \int_G \frac{\mathcal{X}_A}{(1-g)}(1-g)d\nu \\ &= \int_G \frac{\mathcal{X}_A}{(1-g)} g d\mu \\ &= \int_A \underbrace{\frac{\mathcal{X}_G g}{(1-g)}}_{\text{Radon-Nikodym Derivative}} d\mu\end{aligned}$$

That is,

$$\nu(A) = \int_A \frac{\mathcal{X}_G g}{(1-g)} d\mu$$

■

Corollary 16.42.1. *If $\nu(E) = \int_X f d\mu$, then $\nu(X) < \infty \implies f \in L^1(\mu)$.*

Now we're almost ready to tackle the Reisz Representation Theorem for L^p, but we still need some constructions in order to make the proof approachable.

Definition 16.43. *A function ν on a measurable space $(X, \mathscr{A})$ is a signed measure provided:*

1. *For any $E \in \mathscr{A}, \nu(E) \in [-\infty, \infty]$*
2. *$\nu(\emptyset) = 0$*
3. *ν is countably additive.*

Corollary 16.43.1. *If ν is a bounded signed measure and $\nu \ll \mu$ then there exists a function $f \in L^1(\mu)$, not necessarily positive, such that for all $E \in \mathscr{A}$*

$$\nu(E) = \int_E f d\mu$$

Proof. of the **Riesz Representation Theorem**
Case: We assume $\mu(X) < \infty$ and $p > 1$. Let $E \in \mathscr{A}$. Clearly $\mathcal{X}_E \in L^p(X)$. We want to let the operator in question act on this set, and define a signed measure to be this result:

$$\nu(E) := \ell(\mathcal{X}_E)$$

Clearly ν is a bounded signed measure on $(X, \mathscr{A})$. On the other hand, we would like $\nu \ll \mu$. Clearly, when

$$\mu(E) = 0 \implies \ell(\mathcal{X}_E) = \ell(0) = 0 = \nu(E)$$

Therefore, we can use the corollary to the Radon-Nikodym Theorem, providing us a unique function $g \in L^q(\mu)$ such that for all $E \in \mathscr{A} A$

$$\nu(E) = \ell(\mathcal{X}_E) = \int_X \mathcal{X}_E g d\mu$$

To extend this to all of $L^p(X)$, we let

$$\phi = \sum_{n=1}^N a_n \mathcal{X}_{E_n}$$

be a simple function. Therefore,

$$\ell(\phi) = \sum_{n=1}^{N} a_n \ell(\mathcal{X}_{E_n}) = \sum_{n=1}^{N} a_n \int_X \mathcal{X}_{E_n} g d\mu = \int_X \phi g d\mu$$

By the density of the simple functions in L^p, we see

$$\|g\|_{L^q} = \sup_{\|\phi\|_{L^p}=1, \phi \text{ simple}} \int_X \phi \cdot g d\mu = \ell(\phi) \leq \|\ell\| \, \|\phi\|_{L^p} = \|\ell\| < \infty$$

Therefore, $g \in L^q(X)$.

Lastly, take $\{\phi_n\}_n$ a sequence of increasing simple functions such that

$$\phi_n \to f \text{ in } L^p$$

So observe:

$$|\ell(f) - \ell(\phi_n)| \leq \|\ell\| \, \|f - \phi_n\|_{L^p}$$

$$\implies \lim_{n\to\infty} |\ell(f) - \ell(\phi_n)| = 0$$

$$\implies \ell(f) = \lim_{n\to\infty} \ell(\phi_n) = \lim_{n\to\infty} \int_X \phi_n g d\mu$$

We can see by Hölder's inequality:

$$|\phi_n g| \leq 2|f| \cdot g \in L^1$$

Also, $\phi_n g \to f \cdot g$ almost everywhere. Therefore, by Dominated Convergence Theorem, we arrive at

$$\ell(f) = \int_X f \cdot g d\mu$$

Case: Let $p > 1$ but we now we only assume the space is σ-finite. That is ,

$$X = \bigcup_{n=1}^{\infty} X_n$$

with $X_{n+1} \supset X_n$ and $\mu(X_n) < \infty$. Therefore, define the sequence of functions $f_n := f\mathcal{X}_{X_n}$, which has the properties that

- $f_n \to f$ a.e.
- $f_n \in L^p(X_n, \mu)$

Therefore, there must exists a $g_n \in L^q(X_n, \mu)$ such that

$$\ell(f_n) = \int_{X_n} f_n g_n d\mu = \int_X f_n g_n d\mu$$

since $supp(g_n) \subset X_n$.

$$\implies |\ell(f_n)| \leq \|\ell\| \, \|f\|_{L^p} \implies |\int_X f_n g_n d\mu| \leq \|\ell\|$$

So for all n :

$$\implies \|g_n\|_{L^q(X)} \leq \|\ell\|$$

Therefore,

$$\lim_{n\to\infty} g_n(x) = g(x)$$

At this point, we have $f_n \cdot g_n \to f \cdot g$ a.e. and $g \in L^q(X)$. So

$$\ell(f) = \lim_{n\to\infty} \ell(f_n) = \lim_{n\to\infty} \int_X f_n \cdot g_n d\mu$$

Also, $|f_n \cdot g_n| \leq |f||g| \in L^1$ by Hölder because $f \in L^p, g \in L^q$. Therefore, we can apply dominated convergence theorem to conclude

$$\ell(f) = \int_X fg d\mu$$

■

Corollary 16.43.2. *For $1 \leq p < \infty, (L^p(X))^* \cong L^q(X)$. Also, $p = \infty$, then $(L^\infty)^* \supset L^1(X)$.*

Remark 16.44. *For $p = \infty$, the set inequality is necessary.*

Example 16.45. *Consider the measure space $(X, \mathscr{B}(X), \lambda)$ with $X = \mathbb{R}$. Then the set of continuous functions on $\mathbb{R}$ is a closed subspace of bounded functions*

$$C_c(\mathbb{R}) \subset L^\infty(\mathbb{R})$$

Therefore, by Hahn-Banach, we can extend a bounded linear functional defined on $C_c(\mathbb{R})$ to the space $L^\infty(\mathbb{R})$. We want to define ℓ such that $\ell(f) \neq \int f \cdot g dx$ for $g \in L^1(\mathbb{R})$. Define

$$\ell(f) = f(0)$$

We consider the sequence f_n of decreasing triangle functions such that

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0 & x \neq 0 \\ 1 & x = 0 \end{cases}$$

Therefore,

$$\ell(f_n) = f_n(0) = 1$$

But for any $g \in L^1(\mathbb{R})$ we see

$$\int_0^1 f_n g d\mu \to 0$$

Therefore, by extension by Hahn Banach, this property carries to $L^\infty(X)$.

16.6 Weak Convergence

Definition 16.46. *A sequence $X_n \in B$, a Banach space, converges weakly to $x \in B$ provided*

$$\forall f \in B^*, f(x_n) \to f$$

denoted $x_n \rightharpoonup x$.

Example 16.47. *On $L^2([0,1])$, $\sin(nx) \rightharpoonup 0$ in $L^2([0,1])$ as a consequence of Riemann-Lebesgue or Bessel's Inequality.*

Lemma 16.48. *Suppose X is σ-finite with $1 \leq p < \infty$. If $f_n \to f$ in $L^p(X)$, then $f_n \rightharpoonup f$ in $L^p(X)$*

Proof. All we need to show is $\forall g \in L^q$,

$$\int f_n g \to \int fg$$

It suffices to show

$$\int |f_n - f||g| \to 0$$

as $n \to \infty$. We can prove the even stronger condition by Hölder that

$$\int |f_n - f||g| \leq \|f_n - f\|_{L^p} \|g\|_{L^q} \to 0$$

as $n \to \infty$. ■

Theorem 16.49. *Suppose $1 < p < \infty$ and $(X, \mathscr{A}, \mu)$ is a σ-finite measure space. Suppose $f_n \to f$ a.e. and $f_n \in L^p(X)$ with $\|f_n\|_{L^p(X)} \leq B < \infty$. Then $f_n \rightharpoonup f$ in $L^p(X)$.*

Remark 16.50. *Fails for $p = 1$ and $p = \infty$.*

Example 16.51. *On $[0,1], f_n(x) = ne^{-nx} \to 0$ a.e.*

Remark 16.52. *σ-finitedness is necessary.*

Proof. Of Theorem
It suffices to check for all $g \in L^q(X)$ such that

$$\lim_{n\to\infty} \int_X |f_n - f||g| d\mu = 0$$

Clearly, we may assume $B \neq 0$ and by σ-finitedness, we assume

$$X = \bigcup_{n=1} X_n, \quad \mu(X_n) < \infty \quad X_n \subset X_{n+1}$$

Now, for any $g \in L^q(X)$, there exists an $N \in \mathbb{N}$ such that forall $n > N$

$$\int_{X \setminus X_n} |g|^q d\mu < \left(\frac{\epsilon}{3}\right)^q$$

Let $E = X_n \implies \mu(E) < \infty$. We can then select δ such that if $\mu(A) < \delta$,

$$\int_A |g|^q d\mu \leq \left(\frac{\epsilon}{3\max\{1, B\}}\right)^q$$

Applying Egorov's Theorem with respect to this choice of δ, there exists a measurable subset $A \subset E$ such that the $\mu(A) < \delta$ and

$$f_n \rightrightarrows f \text{ uniformly on } E \setminus A$$

Now, we have the tools to tackle the integral in question. Observe, we can split our domain X into three parts,

$$X = A \cup (E \setminus A) \cup (X \setminus E)$$

which split our integral in 3 parts:

$$\int_X |f_n - f||g| d\mu = \underbrace{\int_A |f_n - f||g| d\mu}_{I} + \underbrace{\int_{E\setminus A} |f_n - f||g| d\mu}_{II} + \underbrace{\int_{X \setminus E} |f_n - f||g| d\mu}_{III}$$

Observe,

$$I \le \underbrace{\|f_n - f\|_{L^p}}_{\le 2B} \|g\|_{L^q(A)} \le \frac{2B\epsilon}{3\max\{1, 2B\}} \le \frac{\epsilon}{3}$$

Consider II, we recall that we can select n large enough such that

$$|f_n - f| \le \frac{\frac{\epsilon}{3}}{\|g\|_{L^1(E\setminus A)}}$$

Therefore,

$$II \le \int_{E\setminus A} \frac{\frac{\epsilon}{3}}{\|g\|_{L^1(E\setminus A)}} |g| d\mu = \frac{\epsilon}{3}$$

Lastly,

$$III \le \|f_n - f\|_{L^p(X)} \|g\|_{L^q} \le 2B\frac{\epsilon}{3\max\{1, 2B\}} \le \frac{\epsilon}{3}$$

Therefore, we conclude that

$$\int_X |f_n - f||g| d\mu < \epsilon$$

which gives us weak convergence. ■

17 Differentiation Theorems

Definition 17.1. *Let $f \in L^1_{loc}(\mathbb{R}^n)$ (i.e.*

$$\int_K |f| d\mu < \infty$$

for any compact set $K \subset \mathbb{R}^n$), then we can define the Hardy-Littlewoord Maximal Function by

$$\mu f(x) := \sup_{r>0} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} |f(y)| \, dy$$

Theorem 17.2 (Maximal Function Theorem)**.** *For any $f \in L^p(\mathbb{R}^n)$ with $1 \le p < \infty$, then $\mu f(x) < \infty$ a.e. Morevoer, if $1 < p < \infty$, then*

$$f \in L^p(\mathbb{R}^n) \implies \mu f \in L^p(\mathbb{R}^n)$$

On the other hand, if $p = 1$, then $f \in L^1(\mathbb{R}^n) \implies \mu f \in L^{1,weak}(\mathbb{R}^n)$.

Remark 17.3. *There exists an $f \in L^1(\mathbb{R})$ such that $\mu f \notin L^1$*

Example 17.4. *Consider $f \in C_c^\infty(B_R(0)) \subset L^1(\mathbb{R})$ with $\|f\|_{L^1} \neq 0$. Then*

$$\mu f(x) \ge \frac{1}{\mu(B_{|x|+R}(x))} \int_{B_{|x|+R}(x))} |f(y)| \, dy = \frac{\|f\|_{L^1(\mathbb{R}^n)}}{\omega_n(|x|+R)^n}$$

where $\omega_n = \mu(B_1(0))$ embedded in $\mathbb{R}^n$. Therefore,

$$\int_{\mathbb{R}^n} \mu f(x) \, dx = C \int_0^{2\pi} \int_0^\infty \frac{1}{r^n} r^{n-1} dr d\theta \to \infty$$

Now, in order to prove the Maximal Function Theorem, we need some constructions to help our proof. The first is a Vitali-Covering Lemma:

Lemma 17.5 (5-times Covering Lemma)**.** *Let $\mathcal{B}$ be a collection of balls with positive radii. Further, we suppose*

$$\sup_{B \in \mathcal{B}} \{\text{radius of } B\} = R < \infty$$

Then there exists a subcollection $\mathcal{B}' \subset \mathcal{B}$ such that $\mathcal{B}'$ consists of pairwise disjoint balls and

$$\bigcup_{B \in \mathcal{B}} B \subset \bigcup_{B \in B'} 5B$$

where $5B$ indicates a concentric dilation of factor 5. Further, if the metric space is separable, then $\mathcal{B}'$ can be taken to be countable

$$\bigcup_{B \in \mathcal{B}} B \subset \bigcup_{i=1}^\infty 5B_i \quad B_i \cap B_j = \emptyset$$

Proof. Let $R := \sup_{B \in \mathcal{B}} \{\text{radius of } B\} < \infty$. We want to define a new family of balls :

$$\mathcal{F}_i := \{B \in \mathcal{B} : \frac{R}{2^{i+1}} < \text{ radius of } B \le \frac{R}{2^i}\}$$

for any $i = 0, 1, 2, \ldots$. Now consider $\mathcal{B}_0$ be the maximal collection of balls in F_0 that are pairwise disjoint. Let $\mathcal{B}_1$ be the maximal collection of pairwise disjoint balls in F_1 such that the balls are pairwise disjoint to those of $\mathcal{B}_0$. Continue enumerating maximal collections of pairwise disjoint balls $\{\mathcal{B}_i\}_{i=1}^{n=1}$. Define the collection after this enumeration:

$$\mathcal{F}_n \bigcap \{B \cap B' \neq \emptyset \forall B' \in \bigcup_{i=1}^{n-1} \mathcal{B}_i\}$$

We consider $\mathcal{B}_n$ to be the maximal pairwise disjoint balls serving as such a collection. Further, we define:

$$\mathcal{B}' := \bigcup_{n=0}^{\infty} \mathcal{B}_n$$

Claim: $\bigcup_{B\in\mathcal{B}} \subset \bigcup_{B\in\mathcal{B}'} 5B$
If $B \in \mathcal{B}$, then $B \in \mathcal{F}_k$ for some k. So either $B \in \bigcup_{n=0}^{k} \mathcal{B}_n$ or $B \notin \bigcup_{n=0}^{k} \mathcal{B}_n$.
Case: $B \in \bigcup_{n=0}^{k} \mathcal{B}_n \implies$ conclusion is immediate.
Case: $B \notin \bigcup_{n=0}^{k} \mathcal{B}_n$. Then

$$B \cap B' \neq \emptyset$$

for some $B' \in \bigcup_{n=0}^{k} \mathcal{B}_n$.

- How small is B'? Observe

$$\frac{1}{2^{k+1}} < \text{radius} B'$$

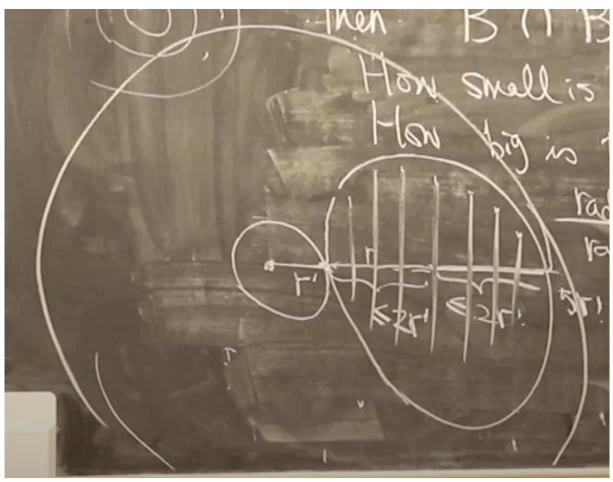
- How big is B? Observe:

$$\text{radius} B \leq \frac{1}{2^k}$$

Therefore, dividing this, we see the following ratio is bounded:

$$\frac{\text{radius } B}{\text{radius } B'} \leq 2 \implies B \subset 5B'$$

Visually:



■

Now we have the tools to tackle on the proof of the Hardy-Littlewoord Maximal Function.

Proof. (Of Hardy-Littlewoord Maximal Function)
Recall: Let the metric space be separable, with

$$\bigcup_{\alpha\in I} B_\alpha \qquad B_\alpha \bigcap B_\beta = \emptyset, \alpha, \beta \in I, \alpha \neq \beta$$

In order for the collection to be dense in the metric space, there must exist a point from a countable dense set in every ball. Therefore, the collection I must also be countable.

Claim: Let $f \in L^1(\mathbb{R}^n)$. We claim $\mu f \in L^{1,weak}(\mathbb{R}^n)$. That is,

$$\lambda(\{x \in \mathbb{R}^n : \mu f(x) > t\}) \leq \frac{C}{t}$$

If $f \in L^1(\mathbb{R}^n) \implies \|f\|_{L^1(\mathbb{R}^n)} < \infty$. Now define the set

$$A_t := \{x \in \mathbb{R}^n : \mu f(x) > t\}$$

For each $x \in A_t$, there must exist an $r_x > 0$ such that

$$\frac{1}{\lambda(B_{r_x}(x))} \int_{B_{r_x}(x)} |f| d\lambda \geq t$$

$$\implies \lambda(B_{r_x}(x)) \leq \frac{1}{t} \int_{B_{r_x}(x)} |f| d\lambda \leq \frac{\|f\|_{L^1(\mathbb{R}^n)}}{t}$$

But since $A_t \subset \bigcup_{x \in \mathbb{R}^n} B_{r_x}(x)$, we take notice that

$$\sup_{x \in \mathbb{R}^n} \{r\} < \infty$$

by the previous inequality. Therefore, we can apply the 5-times covering lemma,

$$\implies A_t \subset \bigcup_{i=1}^{\infty} 5B_{r_i}(x_i) = \bigcup_{i=1}^{\infty} B_{5r_i}(x_i)$$

$$\implies \lambda(A_t) \leq \sum_{i=1}^{\infty} 5^n \lambda(B_{r_i}(x_i)) = \frac{5^n}{t} \int_{\bigcup_{i=1}^{\infty} B_{r_i}(x_i)} |f| dx \leq \frac{5^n \|f\|_{L^1(\mathbb{R}^n)}}{t}$$

On the other hand, if we let $f \in L^\infty$, then it must follow that

$$\mu f = \sup_{r>0} \frac{1}{\lambda(B)} \int_B |f| d\lambda \leq \|f\|_{L^\infty}$$

By the previous claim and note, we see that:

$$\mu|_{L^\infty} : L^\infty \to L^\infty$$

$$\mu|_{L^1} : L^1 \to L^{1,weak}$$

are well-defined and bounded, so by the Marcin Liewicz Interpolation Theorem, we conclude that $\mu : L^p \to L^p$ is a bounded linear operator for all $1 \leq p < \infty$.

■

Theorem 17.6 (Lebesgue Differentiation Theorem)**.** *Suppose $f \in L^1_{loc}(\mathbb{R}^n)$. Then*

$$\lim_{r \to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} f(y)\ dy = f(x)\ a.e.$$

Proof. Define the function:

$$f_r(x) := \frac{1}{|B_r(x)|} \int_{B_r(x)} f(y)\ dy$$

Now, define the functional $\omega : L^1(\mathbb{R}^n) \to \mathbb{R}$ by

$$\omega(f) = \limsup_{r \to 0} f_r(x) - \liminf_{r \to 0}$$

Therefore, if we can prove the following conditions:

1. $\omega(f) = 0$ a.e.
2. $f_r \to f$ in $L^1(\mathbb{R}^n)$

Then (1) $\implies f_r \to g$ a.e. as $r \to 0$, and (2) $\implies \exists$ a subsequence $\{f_{r_i}\}_i \to f$ a.e. Since the subsequence must have the same limit as the original sequence, we can conclude $g = f$ a.e. Therefore, we simply need to prove the conditions before to satisfy our claim:

1. Observe, the following properties of ω:
$$\begin{aligned} \omega(f+g) &\leq \omega(f) + \omega(g) \\ h \text{ continuous } &\implies \omega(h) = 0 \\ \omega(f) &\leq 2\mu f(x) \end{aligned}$$
Claim: $|\{x \in \mathbb{R} : \omega(f) > \epsilon\}| < \epsilon$
If f were continuous, then the above is obvious. However, when $f \in L^1$, we can infinitely approximate f by continuous function h such that
$$\|f - h\|_{L^1} < \frac{\epsilon^2}{2}$$
By the subadditivity of ω,
$$\omega(f) \leq \omega(f-h) + \omega(h) = \omega(f-h)$$
Therefore,
$$\begin{aligned} \{x \in \mathbb{R} : \omega(f) > \epsilon\} &\subset \{x \in \mathbb{R} : \omega(f-h) > \epsilon\} \\ &\subset \{x \in \mathbb{R} : 2\mu(f-h) > \epsilon\} \end{aligned}$$
Therefore,
$$|\{x \in \mathbb{R} : \omega(f) > \epsilon\}| \leq |\{x \in \mathbb{R} : 2\mu(f-h) > \epsilon\}| \leq \frac{2\,\|f-h\|_{L^1}}{\epsilon} < \epsilon$$
2. *Claim:* $f_r \to f$ in L^1 as $r \to 0$.
Observe:
$$\begin{aligned} \int_{\mathbb{R}^n} |f_r(x) - f(x)| dx &= \int_{\mathbb{R}^n} \left| \frac{1}{|B_r(x)|} \int_{B_r(x)} f(y) - f(x) dy \right| dx \\ &= \int_{\mathbb{R}^n} \left| \frac{1}{|B_r(0)|} \int_{B_r(0)} f(x+y) - f(x) dy \right| dx \\ &\leq \frac{1}{|B_r(x)|} \int_{B_r(0)} \left[\int_{\mathbb{R}^n} |f(y+x) - f(x)|\, dx \right] dy \\ &= \lim_{y \to 0} \int_{\mathbb{R}^n} |f(y+x) - f(x)|\, dx = 0 \end{aligned}$$

■

Question 17.7. *Given $f \in L^1(\mathbb{R}^n)$, can we ask how close is f to an $L^\infty(\mathbb{R}^n)$ function?*

As it turns out, we can decompose the function f as the sum of two functions:
$$f = g + b$$
where g is the good function in L^∞ but b is bas lying in L^1. But we can ensure b is incredibly small by stipulating
$$\frac{1}{|Q_i|} \int_{Q_i} b = 0$$
for some collection of cubes $\{Q_i\}$. We will make this more precise in the following theorem:

Theorem 17.8 (Calderon-Zygmund Decomposition)**.** *Given $f \in L^1(\mathbb{R}^n)$ and $\lambda > 0$, there exists $\{Q_i\}_i$ of pairwise disjoint cubes of the form:*

$$Q_i = \prod_{i=1}^{n} [\alpha_i, \alpha_i + \beta).$$

that satisfy the following properties:

1. $|\bigcup_i Q_i| \leq \frac{\|f\|_{L^1}}{\lambda}$
2. *On* $\mathbb{R}^n \setminus (\bigcup_i Q_i)$, $f \leq \lambda$ *a.e.*
3. *For every* $i \in \mathbb{N}$,
$$\lambda < \frac{1}{|Q_i|} \int_{Q_i} f \; dx \leq 2^n \lambda$$

Proof. To prove this, we will use a dyadic decomposition argument.
Suppose $f \in L^1(\mathbb{R}^n), f \geq 0$ and

$$\int_{\mathbb{R}^n} f = L < \infty$$

Let $Q_d(x)$ be a big cube with side length d centered at x such that

$$\frac{1}{|Q_d(x)|} \int_{Q_d(x)} f(y) dy \leq \lambda$$

for all x. This can be achieved by adjusting λ such that

$$\frac{1}{\lambda} \int_{Q_d(x)} f(y) dy \leq |Q_d(x)|$$

for all i. Now tile $\mathbb{R}^n$ by a gird of cubes of this side length d. We will now consider a Dyadic Cutting Procedure:

1. For each Q, we say Q is O.K. if
$$\frac{1}{|Q|} \int_{Q} f \; dy > \lambda$$
2. If Q is NOT O.K., then we will bisect Q into 2^n congruent cubes.
3. For each Subcube of Q, repeat the same procedure starting at 1.

Post condition: After terminating, we will arrive at a collection $\{Q_i\}_i$ of all O.K. cubes, not of the same length anymore. Moreover, we stipulate the following child relation:

$$Q^+ \sim Q \iff \text{If } Q \text{ arises from bisecting } Q^+$$

That is, Q is a child of Q^+. Reversibly, Q^+ is a parent of Q. Now, as a result of our collection, it is clear we satisfy the condition:

$$\begin{aligned} \left| \bigcup_{i=1}^{n} Q_i \right| &= \sum_{i=1}^{n} |Q_i| && Q_i \text{ is O.K.} \\ &< \sum_{i=1}^{n} \frac{1}{\lambda} \int_{Q_i} f dx \\ &\leq \frac{1}{\lambda} \|f\|_{L^1(\mathbb{R}^n)} \end{aligned}$$

That is, we achieve the first condition! Next, for any $x \in \mathbb{R}^n \setminus (\bigcup_i Q_i)$ which are the cubes that are not O.K. from our dyadic procedure, then we see $f(x)$ is dominated by the bound λ. Therefore, consider the cube Q such that $x \in Q$, we see:

$$\frac{1}{|Q|}\int_Q f \leq \lambda$$

Now, letting $|Q| \to 0$, we can conclude by the Lebesgue Differentiation Theorem that

$$f(x) \leq \lambda \ a.e.$$

Lastly, to prove the bounds on the average value, we notice that by the design of our algorithm, we already have the result :

$$\lambda < \frac{1}{|Q_i|}\int_{Q_i} f \ dx$$

On the other hand, if $Q_i \subset Q_i^+$, then we see

$$|Q_i^+| = 2^n|Q_i|$$

Therefore,

$$\begin{aligned}
\frac{1}{|Q_i|}\int_{Q_i} f \ dx &\leq \frac{1}{|Q_i|}\int_{Q_i^+} f \ dx \\
&= \frac{|Q_i^+|}{|Q_i||Q_i^+|}\int_{Q_i^+} f \ dx \\
&= \frac{2^n|Q_i|}{|Q_i||Q_i^+|}\int_{Q_i^+} f \ dx \\
&= \frac{2^n}{|Q_i^+|}\int_{Q_i^+} f \ dx
\end{aligned}$$

■

Question 17.9. *Recall the fact that $f \in L^1 \implies \mu f \in L^{1,weak}$. For what f is $\mu f \in L^1$?*

Theorem 17.10 (Stein)**.** *$\mu f \in L^1(B_R(0)) \iff f \in f \in L\log L(B) \iff \int_{B_R(0)} |f|\log(\max\{|f|,1\})dx < \infty$. That is,*

$$\bigcap_{p>1} L^p \subset L\log L \subset L^1$$

Before we can prove this, we need a stronger Hardy-Littlewood Inequality:

Lemma 17.11.

$$|\{x \in \mathbb{R}^n : \mu f > \lambda\}| \leq \frac{5^n 2}{\lambda}\int_{\{x:|f|>\frac{\lambda}{2}\}} |f|dy$$

Proof. Decompose our function:

$$f = f\mathcal{X}_{\{x:|f|\leq\frac{\lambda}{2}\}} + \underbrace{f\mathcal{X}_{\{x:|f|>\frac{\lambda}{2}\}}}_{f_1}$$

$$\implies \mu f \leq \frac{\lambda}{2} + \mu f_1$$

Therefore, we can apply the Standard Hardy-Littlewood Inequality on f_1 to get the bound:

$$\begin{aligned} |\{x \in \mathbb{R}^n : \mu f > \lambda\}| &\leq |\{x \in \mathbb{R}^n : \mu f_1 > \frac{\lambda}{2}\}| \\ &\leq \frac{5^n}{\lambda} \int_{\mathbb{R}^n} |f_1| dx \\ &= \frac{5^n 2}{\lambda} \int_{|f|>\lambda/2} |f| dy \end{aligned}$$

■

Turns out the midterm was based entirely on this last lecture. So for anyone who thought there would be an equitable distribution of problems across the material taught this quarter can go fuck themselves.

Proof. (Of Stein)
($\Rightarrow$)Calculating the L^1 of μf we see,

$$\begin{aligned} \int \mu f d\mu &= \int_0^\infty |\{x \in B : \mu f > \lambda\}| d\lambda \\ &= \underbrace{\int_0^1 |\{x \in B : \mu f > \lambda\}| d\lambda}_{\leq |B|} + \underbrace{\int_1^\infty |\{x \in B : \mu f > \lambda\}| d\lambda}_{II} \end{aligned}$$

Where

$$\begin{aligned} II &\leq \int_1^\infty \frac{5^n 2}{\lambda} \int_{|f|>\frac{\lambda}{2}} |f| dy \\ &= 5^n 2 \int_B |f| \int_1^{\max\{1,2|f|\}} \frac{1}{\lambda} d\lambda \, dx \\ &= 5^n 2 \int_B |f| \log(\max\{1, 2|f|\}) dx \end{aligned}$$

Therefore,

$$f \in L \log L$$

■

To prove the converse, we need another inequality.

Lemma 17.12 (Reverse Hardy-Littlewoord Maximal Inequality)**.**

$$|\{x \in \mathbb{R}^n : \mu f > 2^{-n}\lambda\} \geq \frac{1}{2^n \lambda} \int_{|f| \geq \lambda} |f| dx$$

Proof. Assuming $f \in L^1(\mathbb{R})$, we can apply the CZ-decomposition, which states there exists the pairwise disjoint sequence $\{Q_i\}_i$

1. $| \cup_i Q_i| \leq \frac{\|f\|_{L^1}}{\lambda}$
2. On $(\cup_i Q_i)^c$, $|f| \leq \lambda$ a.e.
3. $\lambda < \frac{1}{|Q_i|} \int_{Q_i} |f| \leq 2^n \lambda$

Therefore, for every $x \in \bigcup_i Q_i$,

$$\begin{aligned}
\frac{1}{2^n\lambda}\int_{|f|>\lambda} |f|dx &= \frac{1}{2^n\lambda}\sum_i \int |f|dx \\
&\leq \frac{1}{2^n\lambda} 2^n\lambda \sum_i |Q_i| && \text{Apply (3) of CZ-decomposition} \\
&\leq \sum_i |Q_i|
\end{aligned}$$

Claim: $\bigcup_i Q_i \subset \{x \in \mathbb{R}^{n}\text{''}\mu f > 2^{-n}\lambda\}$
Suppose $x \in Q_i$ for some Q_i. Then by property (3) again:

$$\lambda < \frac{1}{|Q_i|}\int_{Q_i} |f| \leq \frac{1}{|Q_i|}\frac{|Q_i^+|}{|Q_i^+|}\int_{Q_i^+} |f| = \frac{2^n}{|Q_i^+|}\int_{Q_i^+} |f| \leq 2^n \mu f(x)$$

$$\implies \mu f(x) > 2^{-n}\lambda$$

Therefore,

$$\frac{1}{2^n\lambda}\int_{|f|>\lambda} |f|dx \leq \sum_i |Q_i| \leq |\{x \in \mathbb{R}^{n}\text{''}\mu f > 2^{-n}\lambda\}|$$

■

Proof. (Continuing the Proof of Stein's theorem)
(⇒) Suppose $\mu f \in L^1(\mathbb{R}^n)$. Then

$$\begin{aligned}
\int |\mu f| &= \int_0^\infty |\{x : \mu f > 2^{-n}\lambda\}| d(2^{-n}\lambda) \\
&\geq 2^{-n}\int_0^\infty \frac{1}{2^n\lambda}\int_{|f|>\lambda} |f| dx d\lambda \\
&\geq \frac{1}{2^{-2n}}\int |f| \int_1^{\max\{1,|f(x)|\}} \frac{1}{\lambda} d\lambda dx \\
&= \frac{1}{2^{2n}}\int_{\mathbb{R}^n} |f| \log(\max\{1, |f(x)|\}) dx < \infty
\end{aligned}$$

■

18 Convolution

Definition 18.1. *If $f, g : \mathbb{R}^n \to \mathbb{R}$, then*

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

Lemma 18.2. *Suppose $f, g \in L^1(\mathbb{R}^n)$. Then for almost every x,*

$$f(x-y)g(y) \in L^1(dy)$$

and also

$$\|f * g\|_{L^1} \leq \|f\|_{L^1} \|g\|_{L^1}$$

Proof. Notice that $f(x-y)g(y)$ is a measurable function on $\mathbb{R}^{2n}$ with respect to the product measure. So by applying Tonelli's theorem, we see:

$$\int |f(x-y)g(y)|\ d(x \otimes y) = \int \int |f(x-y)g(y)| dx\ dy = \int |f(x)| dx \int |g(y)| dy < \infty$$

Since Tonelli's Theorem holds, we can turn to Fubinit to conclude that for almost all x,

$$f(x-y)g(y) \in L^1(dy)$$

On the other hand, we see:

$$\begin{aligned}
\int |f * g|(x) dx &= \int \left| \int f(x-y)g(y) dy \right| dx \\
&\leq \int \int |f(x-y)||g(y)| dx\ dy \\
&= \int |f(x-y)| dy \int |g(y)| dy \\
&= \|f\|_{L^1} \|g\|_{L^1}
\end{aligned}$$

∎

Lemma 18.3 (Convolution is Associative and Commutative)**.** *$f * g = g * f$ and $f * (g * h) = (f * g) * h$. Moreover, the space $(L^1(\mathbb{R}^n), *)$ is a commutative Banach algebra.*

Theorem 18.4. *Suppose $K(x) \in L^1(\mathbb{R}^n)$, in which we will refer to this function as the kernel, and define a map $L : L^1(\mathbb{R}^n) \to L^1(\mathbb{R}^n)$ by*

$$L(f) = K * f$$

Then L is a bounded linear operator with $\|L\| \leq \|K\|_{L^1}$.

Question 18.5. *Can we extend this theorem to L^p. That is, a function $K \in L^p(\mathbb{R}^n)$ such that*

$$f \in L^p(\mathbb{R}^n) \implies \|K * f\|_{L^p(\mathbb{R}^n)} < \infty$$

Remark 18.6. *Suppose $f \in L^\infty(\mathbb{R}^n)$. Then*

$$\begin{aligned}
\|K * f\|_{L^\infty} &= \left\| \int K(x-y)f(y)\ dy \right\|_{L^\infty} \\
&\leq \|f\|_{L^\infty} \|K\|_{L^1}
\end{aligned}$$

Therefore, $K* : L^1 \to L^1$ is a bounded linear operator, and $K* : L^\infty \to L^\infty$ is a bounded linear operator. Therefore, by MarcinKiewicz Interpolation Theoem, we conclude the following theorem:

Theorem 18.7. *If $K \in L^1$, and $1 \leq p \leq \infty$, then*

$$\|K * f\|_{L^p} \leq C \|f\|_{L^p}$$

Example 18.8. *Let $r > 0$ and define the kernel:*

$$K_r(x) := \frac{1}{|B_r(0)|}\mathcal{X}_{B_r(0)}$$

Notice $\|K_r\|_{L_1} = 1$. So,

$$(K_r * f)(x) := \frac{1}{|B_r(0)|}\int_{\mathbb{R}^n} \mathcal{X}_{B_r(0)}(x-y)f(y)\ dy = \frac{1}{|B_r(0)|}\int_{B_r(x)} f(y)dy$$

Therefore, this kernel gives us the average value of f over the ball of radius r centered at x. Clearly this is bounded by the maximal function. Therefore,

$$\|K_r * f\|_{L^p} \leq C \|f\|_{L^p}$$

18.1 Motivating Approximation Identities

Recall we can define the translation $(\tau_y f)(x) = f(x-y)$ and the following facts about this translation:

- $\|\tau_y f\|_{L^p(\mathbb{R}^n)} = \|f\|_{L^p(\mathbb{R}^n)}$
- $\lim_{y\to 0} \|\tau_y f - f\|_{L^p(\mathbb{R}^n)} = 0$

Upon convolving with f, we see:

$$(K * f)(x) = (f * K) = \int f(x-y)K(y)\ dy = \int \tau_y f(x) K(y) dy$$

We will use these facts in order to introduce an intriguing class of approximations.

If $K \in L^1$, with

$$K_r(x) = \frac{1}{r^n} K\left(\frac{x}{r}\right)$$

Then

$$\int K_r(x)\ dx = \int K\left(\frac{x}{r}\right)\frac{dx}{r^n} = \int K(y)dy$$

Now suppose for some $\delta > 0$

$$\int_{|x|>\delta} |K_r(x)| dx = \int_{|x|>\delta} \frac{1}{r^n}\left|K\left(\frac{x}{r}\right)\right|\ dx = \int_{|y|>\frac{\delta}{r}} K(y)\ dy$$

Notice the integrand is now independent of r. Therefore,

$$\int_{|y|>\frac{\delta}{r}} K(y)\ dy \to 0 \text{ as } r \to \infty$$

Therefore,

$$\boxed{\lim_{r\to 0}\int_{|x|>\delta} |K_r(x)|\ dx = 0}$$

This property is very desirable. Another very desirable property is

$$\boxed{\int K(x)\ dx = 1}$$

If such a kernel K were to possess both properties, then the following definition applies:

Definition 18.9. *Suppose $K \in L^1$ satisfies the following properties:*

1. $\lim_{r\to 0} \int_{|x|>\delta} |K_r(x)|\ dx = 0$
2. $\int K(x)\ dx = 1$

Then K is called an approximation identity.

Question 18.10. *So where does the term identity come into play?*

Normally, when we have an identity, we simply multiply them to get whatever the input function was. But firstly, these approximation identities only return an approximation upon taking an operation with an original function. Moreover, this operation is no longer multiplication, but instead convolution.
Before we prove that we can take a sequence of convolutions that arbitrarily approach a given function, we need a small mechanism from the following lemma:

Lemma 18.11 (Minkowski's Integral Inequality)**.** *Suppose μ, ν are σ-finite measures on X and Y, and $F : X \times Y \to \mathbb{R}$ a measurable function with respect to $d\mu \otimes d\nu$ Let $1 \le p < \infty$. Then*

$$\left\| \int_X F(x,y) d\mu_x \right\|_{L^p(Y)} \le \int_X \|F(x,y)\|_{L^p(Y)}\, d\mu_x$$

If we replace the integral with a summation, this will appear identical to Minkowski's inequality.

Proof.

$$\begin{aligned}
\left\| \int_X |F(x,y)| d\mu_x \right\|_{L^p(Y)} &= \sup_{\|f\|_{L^q(Y,d\nu)}=1} \int_Y \left[f(y) \int_X |F(x,y)| d\mu_x \right] d\nu_y && \text{Duality of } L^p(Y) \\
&= \sup_{\|f\|_{L^q(Y,d\nu)}=1} \int_X \int_Y |f(y)||F(x,y)|\ d\nu_y\ d\mu_x \\
&\le \sup_{\|f\|_{L^q(Y,d\nu)}=1} \int_X \|f\|_{L^q} \|F(x,\cdot)\|_{L^p(Y,d\nu)}\, d\mu_x && \text{Hölder} \\
&= \int_X \left[\int_Y |F(x,y)|^p d\nu_y \right]^{1/p} d\mu_x
\end{aligned}$$

■

Now we have the necessary tools to tackle one of the main problems of convolution.

Theorem 18.12. *Suppose $K \in L^1(\mathbb{R}^n)$ and*

1. $\int_{\mathbb{R}^n} K(x) = 1$
2. $\lim_{r\downarrow 0} \int_{|x|>\delta} K_r(x) dx = 0$, *where* $K_r(x) = \frac{1}{r^n} K(x/r)$ *a rescaled version of the kernel K.*

Then for $1 \le p < \infty$,

$$\lim_{r\downarrow 0} \|K_r * f - f\|_{L^p} = 0$$

Proof. By the first property,

$$\int_{\mathbb{R}^n} K(x) = 1 \implies \int_{\mathbb{R}^n} K_r(x) = 1$$

Given $\epsilon > 0$, we can find δ such that for any $|y| < \delta$,

$$\|\tau_y f - f\|_{L^p} < \epsilon$$

Then

$$\begin{aligned}
\|K_r * f - f\|_{L^p} &= \left\| \int_{\mathbb{R}^n} K_r(y)\tau_y f(x)dy - f(x) \right\|_{L^p(X)} \\
&= \left\| \int_{\mathbb{R}^n} K_r(y)[\tau_y f(x) - f(x)]dy \right\|_{L^p(X)} && \int_{\mathbb{R}^n} K_r(x) = 1 \\
&\leq \int_{\mathbb{R}^n} |K_r(y)| \, \|\tau_y f(x) - f(x)\|_{L^p(X)} \, dy && \text{Minkowski} \\
&= \underbrace{\int_{|y|>\delta} |K_r(y)| \, \|\tau_y f(x) - f(x)\|_{L^p(X)} \, dy}_{I} + \underbrace{\int_{|y|\leq\delta} |K_r(y)| \, \|\tau_y f(x) - f(x)\|_{L^p(X)} \, dy}_{II}
\end{aligned}$$

By design, when $|y| \leq \delta \implies \|\tau_y f - f\|_{L^p} < \epsilon$, therefore,

$$II = \int_{|y|\leq\delta} |K_r(y)| \, \|\tau_y f(x) - f(x)\|_{L^p(X)} \, dy < \epsilon \int_{|y|\leq\delta} |K_r(y)| \leq \epsilon \int_{\mathbb{R}^n} |K_r(y)| = \epsilon \to 0$$

since this is true for any ϵ. On the other hand, consider the first integral,

$$I \leq 2 \, \|f\|_{L^p} \int_{|y|>\delta} |K_r(y)| dy \to 0$$

by the second assumed property. ∎

Therefore, if you convolve with an approximation identity and take the limit, you will recover your function! To continue, we consider an example.

Example 18.13. *We can construct a $C_c^\infty(\mathbb{R}^n)$:*

$$\phi(x) = \frac{1}{N} e^{\frac{-1}{|x|^2 - 1}}$$

where N is a normalization constant.

Question 18.14. *Why is such a function like this important?*

Corollary 18.14.1 (Smooth Approximations)**.** *Let $f \in L^p(\mathbb{R}^n)$. If $\phi \in C_c^\infty(\mathbb{R}^n)$ and $\phi \geq 0$ and*

1. *$\phi \geq 0$,*
2. *$\int \phi(y)dy = 1$*

Then $\phi_r * f \in C^\infty$ and

$$\lim_{r \downarrow 0} \|\phi_r * f - f\|_{L^p} = 0,$$

where $\phi_r(x)$ is the rescaled version of ϕ defined $\phi_r(x) = \frac{1}{r^n}\phi(x/r)$

Proof. Thanks to continuity and the assumptions on ϕ, it must follow that ϕ_r satisfy the assumptions in the previous theorem, and therefore we immediately arrive at

$$\lim_{r\downarrow 0} \|\phi_r * f - f\|_{L^p} = 0$$

But we still need to prove that $\phi_r * f \in C^\infty$. Looking at the definition of our rescaled kernel:

$$(\phi * f)(x) = \int_r \phi_r(x-y)f(y)dy$$

is an integral and therefore can be differentiated. ■

18.2 Interpolation under Convolutions

Theorem 18.15 (Young's Inequality)**.** *If $f \in L^p(\mathbb{R}^n), g \in L^q(\mathbb{R}^n)$ with $1 \leq p, q \leq \infty$, then*

$$\|f * g\| \leq \|f\|_{L^p} \|g\|_{L^q}$$

with

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

Lemma 18.16. *Suppose $K(x,y)$ is a measurable function on $\mathbb{R}^n \times \mathbb{R}^n$ and*

$$\|K(x,y)\|_{L^q_x L^\infty_y} = \sup_y \left(\int_X |K(x,y)|^q \, dx \right)^{\frac{1}{q}} \leq C$$

for $1 \leq q \leq \infty$. Then we can define

$$Tf(x) = \int K(x,y)f(y)dy$$

Then $1 \leq p \leq r \leq \infty$ in which $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Then

$$\|Tf\|_{L^r} \leq C_q \|f\|_{L^q}$$

Proof. (Of Young's Inequality using Lemma)
Suppose $K(x,y) = g(x-y)$ with $g \in L^q$. Then

$$Tf = \int g(x-y)f(y)dy$$

Now, we let

$$C = \left(\int |g(x-y)|^q \, dx \right)^{1/q} = \|g\|_{L^q}$$

So we see:

$$\|Tf\| = \|f * g\|_{L^r} \leq \|g\|_{L^q} \|f\|_{L^p}$$

■

Proof. (Of Lemma)

$$\begin{aligned} |Tf(x)| &\leq \int |K(x,y)||f(y)|dy \\ &\leq \|K(x,y)\|_{L^q_y L^\infty_x} \|f\|_{L^{q'}} && \frac{1}{q} + \frac{1}{q'} = 1 \\ &\leq C \|f\|_{L^{q'}} \end{aligned}$$

Therefore T is bounded from (q', ∞). We can apply the Reisz-Thorin interpolation theorem by letting $p_1 = q'$ and $p_2 = \infty$. On the other hand,

$$\begin{aligned}\|Tf\|_{L^q} &\leq \left\| \int |K(x,y)||f(y)|dy \right\|_{L^q} \\ &\leq \int \|K(x,y)\|_{L^q_x} |f(y)|dy \\ &\leq C\|f\|_{L^1}\end{aligned}$$

So we let $1 = p_2$ and $q = q_2$. Therefore, Riesz-Thorin Interpolation says

$$\frac{1}{p_\lambda} = \frac{1-\lambda}{p_1} + \frac{\lambda}{p_2}$$
$$\frac{1}{q_\lambda} = \frac{1-\lambda}{q_1} + \frac{\lambda}{q_2}$$

Therefore,

$$\|Tf\|_{L^{q_\lambda}} \leq M_1^{1-\lambda} M_2^{1-\lambda} \|f\|_{L^{p_\lambda}}$$

where $\lambda \in [0,1]$. Further, subbing in our chosen values of p_1, p_2, q_1, q_2:

$$\frac{1}{r} = \frac{1-\lambda}{\infty} + \frac{\lambda}{q}$$
$$\frac{1}{p} = \frac{1-\lambda}{q'} + \frac{\lambda}{1} = (1-\lambda)(1-\frac{1}{q}) + \lambda = 1 + \frac{(\lambda-1)}{q}$$

Therefore,

$$1 + \frac{1}{r} = 1 + \frac{\lambda}{q} = \underbrace{1 + \frac{\lambda}{q} - \frac{1}{q}}_{1/p} + \frac{1}{q} = \frac{1}{p} + \frac{1}{q}$$

■

18.3 Proving the Riesz-Thorin Interpolation Theorem

Recall the Riesz-Thorin Interpolation Theorem:

Theorem 18.17 (Riesz-Thorin Interpolation Theorem)**.** *Suppose we are given two $(X, \mathscr{A}_1, \mu_1), (Y, \mathscr{A}_2, \mu_2)$ σ-finite measure spaces. Also, suppose*

$$1 \leq p_1 \leq p_2 \leq \infty$$
$$1 \leq q_1 \leq q_2 \leq \infty$$

Lastly, let

$$T : L^{p_1}(X) + L^{p_2}(X) \to L^{q_1}(Y) + L^{q_2}(Y)$$

be a linear operator such that the restrictions $T|_{L^{p_1}(X)} : L^{p_1}(X) \to L^{q_1}(Y)$, $T|_{L^{p_2}(X)} : L^{p_2}(X) \to L^{q_2}(Y)$ are well-defined and bounded. Then T is a bounded linear operator

$$T : L^{p_\lambda}(X) \to L^{q_\lambda}(Y)$$

where $\lambda \in [0,1]$ and

$$\frac{1}{p_\lambda} = \frac{1-\lambda}{p_1} + \frac{\lambda}{p_2}$$
$$\frac{1}{q_\lambda} = \frac{1-\lambda}{q_1} + \frac{\lambda}{q_2}$$

Moreover, the norm can be controlled as:

$$\|T\|_{L^{p_\lambda} \to L^{q_\lambda}(Y)} \leq \|T\|^{1-\lambda}_{L^{p_1} \to L^{q_1}} \|T\|^{\lambda}_{L^{p_2} \to L^{q_2}}$$

Proof. Case: $p < \infty$ and $q > 1$.
Define the analytic extensions

$$\begin{cases} \frac{1}{p(z)} &= \frac{1-z}{p_1} + \frac{z}{p_2} \\ \frac{1}{q'(z)} &= \frac{1-z}{q_1'} + \frac{z}{q_2'} \end{cases}$$

where

$$\frac{1}{q_1'} + \frac{1}{q_1} = 1 \quad \frac{1}{q_2'} + \frac{1}{q_2} = 1$$

Therefore, we notice

$$p(0) = p_1, p(1) = p_2 \quad q'(0) = q_1', q'(1) = q_2$$

So our functions $p(z), q(z)$ are in fact analytic extensions. Now it suffices to show:

$$\|T\|_{L^p \to L^q} = \sup_{\|f\|_{L^p} \leq 1, \|g\|_{L^{q'}}} \int (Tf) \cdot g d\nu \leq \|T\|_{L^{p_1} \to L^{q_1}}^{1-\lambda} \|T\|_{L^{p_2} \to L^{q_2}}^{\lambda}$$

To do so, we shall define on the strip $0 \leq Real[z] \leq 1$:

$$f_z(x) = |f(x)|^{\frac{p}{p(z)}} e^{i \cdot arg\{f(x)\}}$$

$$g_z(x) = |g(x)|^{\frac{p}{p(z)}} e^{i \cdot arg\{g(x)\}}$$

As a result, when we evaluate z at θ:

$$f_\theta(x) = f(x) \quad g_\theta(x) = g(x)$$

So these are simply analytic extensions that will help us to our conclusion. Define:

$$F_z(x) := \int (Tf_z(x)) g_z(x) \, d\nu$$

We need to now check the following in order to apply the 3 line lemma to F_z:

1. *Claim:* F_z is bounded and continuous in $0 \leq Real[z] \leq 1$
   Notice, when the operator hits this function in L^p, we have no clue if this lies in L^q or not. So to show this is bounded, we can use a density argument by approximation by simple functions. So because $p < \infty$ and $q > 1$, we can assume f, g are simples functions whose sets are pairwise disjoint:

   $$f = \sum_{i=1}^{F} f_i \mathcal{X}_{F_i} \qquad g = \sum_{j=1}^{G} g_i \mathcal{X}_{G_i}$$

   Therefore,

   $$f_z(x) = \sum_{i=1}^{F} |f_i|^{\frac{p}{p(z)}} e^{i \cdot arg\{f_i\}} \mathcal{X}_{F_i} \qquad g_z(x) = \sum_{i=1}^{G} |g_i|^{\frac{q'}{q'(z)}} e^{i \cdot arg\{g_i\}} \mathcal{X}_{G_i}$$

   Since F_z is a linear operator, then it simply acts upon the indicator functions of its argument, so the convolution becomes:

   $$F_z(x) = \sum_{i,j} |f_i|^{\frac{p}{p(z)}} |g_j|^{\frac{q'}{q'(z)}} e^{i \cdot arg(f_i + g_j)} \int_{G_j} T(\mathcal{X}_{F_i}) d\nu$$

   Notice, $\int_{G_i} T(\mathcal{X}_{F_i})$ is bounded in addition to $|f_i|^{\frac{p}{p(z)}}, |g_j|^{\frac{q'}{q'(z)}}$ since for any real number $\gamma, |\gamma^z| = \gamma^{Re[z]}$ and since $0 \leq Re[z] \leq 1 \implies 0 \leq \frac{1}{p(z)}, \frac{1}{q(z)} \leq 1$. Therefore, $F_z(x)$ is bounded and therefore must also be continuous!

2. *Claim:* F_z is analytic $0 < Real[z] < 1$
   There are no singularities within the strip and the function is written without use of nonholomorphic functions.

3. *Claim:* For all $y \in \mathbb{R}$, $|F_{0+iy}(x)| \leq M_0 = \|T\|^{1-\lambda}_{L^{p_1} \to L^{q_1}}$
   Now, returning to the original formula of our function after having proven boundedness and analyticity, we can apply Hölder's inequality:

$$\begin{aligned} |F_{iy}(X)| &= \left| \int (Tf_{iy}(g_{iy}) \right| \\ &\leq \|Tf_{iy}\|_{L^q} \|g_{iy}\|_{L^{q'}} && \text{Hölder} \\ &\leq \|T\|_{L^{p_1} \to L^{q_1}} \|f\|_{L^{p_1}} \|g\|_{L^{q_1'}} \\ &\leq M_0 \|f\|_{L^{p_1}} \|g\|_{L^{q_1'}} \end{aligned}$$

   So we now observe, by definition

$$\begin{aligned} \|f_{iy}\|^{p_1}_{L^{p_1}} &= \left\| |f(x)|^{\frac{p}{p(iy)}} \right\|^{p_1}_{L^{p_1}} \\ &= \left\| |f|^{Re\left[\frac{p}{p(iy)}\right]} \right\|^{p_1}_{L^{p_1}} \\ &= \left\| |f|^{\frac{p}{p_1}} \right\|^{p_1}_{L^{p_1}} && \text{since } Re\left[\frac{1}{p(iy)} = \frac{1-iy}{p_1} + \frac{iy}{p_1}\right] = \frac{1}{p_1} \\ &= \int |f|^p \leq 1 \end{aligned}$$

   And similarly,

$$\begin{aligned} \|g_{iy}\|^{q_1'}_{L^{q_1'}} &= \left\| |g(x)|^{\frac{q'}{q'(iy)}} \right\|^{q_1'}_{L^{q_1'}} \\ &= \left\| |g|^{\frac{q'}{q_1'}} \right\|^{q_1'}_{L^{q_1'}} \\ &= \int |g|^{q'} \leq 1 \end{aligned}$$

   Therefore,

$$|F_{iy}(X)| \leq M_0 \|f\|_{L^{p_1}} \|g\|_{L^{q_1'}} = M_0$$

4. *Claim:* For all $y \in \mathbb{R}$, $|F_{1+iy}(x)| \leq M_1 = \|T\|^{1-\lambda}_{L^{p_2} \to L^{q_2}}$
   The proof for this is nearly identical to the previous claim!

Because F_z now satisfies the 3-lines lemma, we conclude:

$$|F_z| \leq M_0^{1-\lambda} M_1^{\lambda}$$

where $\lambda = Re[z]$. ■

Note 18.18. *I'm gonna note that it's really difficult to follow Kevin's lectures. He jumps around in logic **a lot** (some might go so far as to say a fuck ton), so please notice how these notes are the Marie Kondo reorganized version of his lectures.*

19 Sobolev Spaces

Definition 19.1. *Multi-Indices: a vector $\alpha \in (\mathbb{Z}_{\geq 0})^n$, which takes the form:*

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$$

with $|\alpha| = \sum_{i=1}^n |\alpha_i|$.

Definition 19.2. *Directional Derivative: If $\phi \in C^\infty(\mathbb{R}^n)$, then*

$$D^\alpha \phi = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} \phi$$

Definition 19.3. *If $u \in C_c^\infty(\mathbb{R}^n)$, then the weak-$\alpha$-derivative of u, denoted $D^\alpha u$, is a $L^1_{loc}(\mathbb{R}^n)$ function v such that*

$$\int v\phi dx = (-1)^{|\alpha|} \int u D^\alpha \phi dx$$

for every $\phi \in C_c^\infty(\mathbb{R}^n)$.

Remark 19.4. *This definition lacks computational relevance. It provides simply a check for a function that you supply.*

Lemma 19.5. *The weak-α-derivative of u is unique.*

Proof. Suppose

$$\int v_1 \phi dx = (-1)^{|\alpha|} \int u D^\alpha \phi dx = \int v_2 \phi dx$$

for any $\phi \in C_c^\infty(\mathbb{R}^n)$.

$$\implies \int (v_1 - v_2)\phi dx = 0$$

Taking a sequence of continuous functions of compact support that approximates $v_1 - v_2$, we realize $v_1 - v_2 = 0$ a.e. Moreover, $v_1 = v_2$ a.e. ∎

Definition 19.6. *Given $\Omega \subset \mathbb{R}^n$ open, with $1 \leq p \leq \infty$ and $K \in \mathbb{Z}_{\geq 0}$, then*

$$W^{k,p}(\Omega) := \{u : \Omega \to \mathbb{R}^n : D^\alpha u \text{ exists } \forall |\alpha| \leq k \text{ and } D^\alpha u \in L^p(\mathbb{R}^n)\ \forall |\alpha| \leq k\}$$

Further, we equip this space with the norm:

$$\|f\|_{W^{k,p}(\Omega)} = \sum_{|\alpha| \leq k} \|D^\alpha f\|_{L^p(\Omega)}$$

Lemma 19.7. *$(W^{k,p}(\Omega), \|\cdot\|_{W^{k,p}(\Omega)})$ is a Banach Space.*

Proof. Suppose $\{u_m\} \subset W^{k,p}(\Omega)$ be a Cauchy sequence. Clearly, this space inherits the the convergence in L^p. That is,

$$u_m \to u \in L^p(\Omega)$$

But now we need to ask if $u \in W^{k,p}(\Omega)$? That is,

$$D^\alpha u_m \to D^\alpha u \in L^p(\Omega)$$

Define $u_\alpha := \lim D^\alpha u_m$. As a consequence of the Dominated Convergence Theorem:

$$\int_\Omega u D^\alpha \phi = \lim_{m \to \infty} \int_\Omega u_m D^\alpha \phi = \lim_{m \to \infty} (-1)^{|\alpha|} \int_\Omega D^\alpha u_m \phi = (-1)^{|\alpha|} \int_\Omega u_\alpha \phi$$

By the uniqueness of the weak-derivative, we conclude $D^\alpha u = u_\alpha$, completing the proof! ∎

Let's build some intuition about weak-derivatives:

Example 19.8. *Consider the function $u : (0, \infty) \to \mathbb{R}$ defined by*

$$u(x) = \begin{cases} x & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

We claim $Du = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & x > 1 \end{cases}$. Observe:

$$\int_0^\infty uD\phi = \int_0^1 x\phi' + \int_1^\infty \phi' = x\phi|_0^1 - \int_0^1 \phi - \phi(1) = \phi(1) - \int_0^1 \phi - \phi_1 = -\int_0^\infty \mathcal{X}_{[0,1]}\phi$$

In essence, weak derivatives do not care about corners. Moreover, weakly-differentiable implies Lipschitz continuity.

19.1 Interior Approximations

Example 19.9. *Consider the space $W^{1,1}(-1,1)$ and consider the function $u \equiv 1$ on $(-1,1)$. We claim $C_c^\infty(-1,1)$ is not dense in $W^{1,1}(-1,1)$ by showing there cannot be a sequence of functions that densely approaches u.*

Suppose $u_m \in C_c^\infty(-1,1)$ of compact support $supp(u_m) \subset (-1,1)$ such that

$$u_m \to u \text{ in } W^{1,1}$$

Due to the compact support of u_m, we know

$$u_m(-1) = u_m(1) = 0 \forall m$$

On the other hand, since these functions have to approximate u, there must be a point x_m for each u_m such that

$$u_m(x_m) \geq \frac{1}{2}$$

Now, exploiting the first-derivative in this Sobolev space:

$$\int_{-1}^1 |u_m'| \geq \left|\int_{-1}^{x_m} u'\right| + \left|\int_{x_m}^1 u'\right| \geq \frac{1}{2} + \frac{1}{2} = 1$$

Therefore, but this can't occur since $u' = 0$ and $u_m' \not\to u'$ in L^p. Therefore, no such sequence can exist.

This example defeats some of our intuition held from L^p spaces, specifically in regards to the density of useful functions like $C_c^\infty(\mathbb{R}^n)$.

Therefore, care must be given in order to find dense sets of continuous functions.

Lemma 19.10 (Interior Approximation)**.** *Let $1 \leq p < \infty$, $\Omega \subset \mathbb{R}^n$ an open region. Given $u \in W^{k,p}(\Omega)$ and ϵ,, we define the set*

$$\Omega_\epsilon := \{x \in \Omega : dist(x, \partial\Omega) > \epsilon\}$$

Then, given $\delta > 0$, there exists $u^\delta \in C^\infty(\mathbb{R}^n) \cap W^{k,p}(\Omega_\epsilon)$ such that

$$\left\|u^\delta - u\right\|_{W^{k,p}(\Omega_\epsilon)} < \delta$$

In other words, this lemma tells us we can get arbitrarily close to a function in our Sobolev space provided we respect the boundary strip of length ϵ.

Proof. Idea: Use convolution to smooth out our $W^{k,p}$ function. But this time, there are more derivatives in play.

Since convolution is effectively a weight average over a small ball, we can ensure our average pertains to our function by staying an ϵ away from the boundary, and then integrating over balls of radius ϵ.

Let $\phi_y(x) \in C_c^\infty(B_\eta(0))$ such that

$$\int \phi_y = 1$$

Therefore, ϕ_y is an approximation identity. Now pick $\eta < \epsilon$. Then by what was previously stated about maintaining distance away from the boundary, we conclude $\phi_y * u$ is well-defined on Ω_ϵ. We need to check that the Derivative converges to the desire function :

$$\begin{aligned} D^\alpha(\phi_\eta * u) &= \int D_x^\alpha \phi_\eta(x-y)u(y)dy \\ &= (-1)^{|\alpha|} \int D_y^\alpha \phi_\eta(x-y)u(y)\ dy \\ &= \int \phi_\eta(x-y) D_y^\alpha u(y)\ dy \\ &= \phi_\eta * \underbrace{D^\alpha u}_{\in L^p} \end{aligned}$$

Therefore, by assumption, is must follow

$$D^\alpha(\phi_\eta * u) = \phi_\eta * D^\alpha u \to D^\alpha u$$

as $\eta \to 0$. Further,

$$\lim_{\eta \downarrow 0} \|\phi_\eta * u - u\|_{W^{k,p}(\Omega_\epsilon)} = 0$$

∎

19.2 Global Approximations

Lemma 19.11 (Meyer-Serrin)**.** *Let $\Omega \subset \mathbb{R}^n$ be an open subset. Let $1 \leq p < \infty$. Given $u \in W^{k,p}(\Omega)$ and $\delta > 0$, there exists $u^\delta \in C^\infty(\Omega) \cap W^{k,p}(\Omega)$ such that*

$$\left\| u^\delta - u \right\|_{W^{k,p}(\Omega)} < \delta$$

Proof. Similar to the proof of the interior approximation, we define a sequence of sets

$$\Omega_0 := \emptyset$$

$$\Omega_j := \{x \in \Omega : dist(x, \partial\Omega) > \frac{1}{j}\}$$

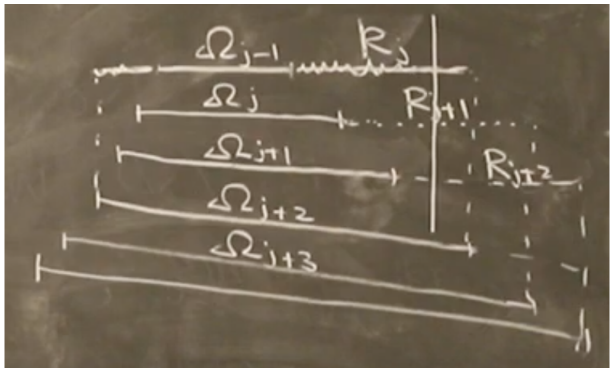
for any $1 \leq j$. Some obvious observations are

$$\Omega_j \subset \Omega_{j+1} \qquad \text{and} \qquad \Omega = \bigcup_{j=0}^\infty \Omega_j$$

Now define $R_j := \Omega_{j+2} \setminus \overline{\Omega_{j+1}}$ to be the closed remainders between successive sets. Therefore ,

$$\Omega = \bigcup_{j=1}^{\infty} R_j$$

We also take note that because of the choose of indexing:



Then we see our remainder sets R_j have the property:

$$\#\{i : R_i \cap R_j \neq \emptyset\} \leq 3$$

Therefore, this yields the property that we can identify a partition of unity $\{\eta_j\}_{j=1}^{\infty} \subset C_c^{\infty}(\mathbb{R}^n)$ such that

$$\textit{supp}\{\eta_j\} \subset R_j \qquad \text{and} \qquad 0 \leq \eta_1 \leq 1$$

and to satisfy unity

$$\sum_{j=1}^{\infty} \eta_j = 1 \quad \text{ on } \Omega$$

Therefore, we can write u using our partition of unity by multiplying both sides by u:

$$u = \sum_{i=1}^{\infty} \eta_i u$$

We can now leverage an interior estimate on each $\eta_i u$. Specifically:

$$\phi_{\epsilon_j} * (\eta_i u) \in C_c^{\infty}(\Omega)$$

where ϕ_ϵ is the smooth approximate identity. Therefore,

$$\left\|\phi_{\epsilon_j} * (\eta_i u) - (\eta_i u)\right\|_{W^{k,p}(\Omega)} < \frac{\delta}{2^j}$$

Summing these interior approxmations as V:

$$u = \sum_{i=1}^{\infty} \eta_i u \implies V(x) = \sum_{j=1}^{\infty} \phi_{\epsilon_j} * (\eta_i u)$$

we see:

$$\|v - u\|_{W^{k,p}(\Omega)} \leq \sum_{j=1}^{\infty} \left\|\phi_{\epsilon_j} * (\eta_i u) - (\eta_i u)\right\|_{W^{k,p}(\Omega)} \leq \sum_{j=1}^{\infty} \frac{\delta}{2^j} = \delta$$

■

Corollary 19.11.1. *If $\Omega' \subset \Omega$ is compactly contained in Ω, that is ,*

$$\Omega' \subset \overline{\Omega'} \textit{ compact} \subset \Omega$$

then $u \in W^{k,p}(\Omega')$ can be arbitrarily approximated by $C_c^{\infty}(\Omega)$.

19.3 Smooth Approximations up to Boundary

Question 19.12. *Can we push these approximations up to the boundary and maintain proximity?*

It turns out, we need to assume some regularity on the boundary in addition to the function spaces.

Definition 19.13. *Given a region Ω, a C^1-boundary $\partial\Omega$ is one such that for every $x_0 \in \partial\Omega$, there exists $r > 0$ and an C^1-embedding-function*

$$\gamma : \mathbb{R}^{n-1} \to \mathbb{R}^n$$

such that

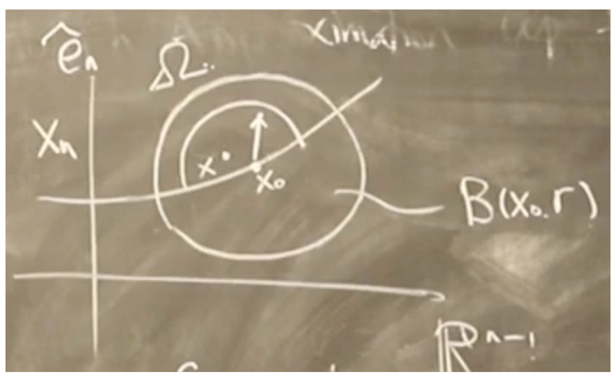
$$B_r(x_0) \cap \Omega = \{(x_1, \ldots, x_n) \in B_r(x_0) : x_n > \gamma(x_1, \ldots, x_{n-1})\}$$

Lemma 19.14. *Let $\Omega \subset \mathbb{R}^n$ open, bounded with C^1 boundary. If $u \in W^{k,p}(\Omega)$, then for any given $\delta > 0$, there exists $u^\delta \in C^\infty(\overline{\Omega})$ and*

$$\left\| u^\delta - u \right\|_{W^{k,p}(\Omega)} < \delta$$

Turns out, we can relax this lemma to include Lipschitz boundaries that possess cusps, but this requires a more involved and unrewarding argument. The only reason we assume C^1 boundary is because there is a clearly defined normal to the boundary $\partial\Omega$ for any value $x_0 \in \partial\Omega$. Otherwise, you could easily edit the proof to take local average of normals which tends to be the generalization for Lipschitz boundaries.

Proof. Suppose $x \in B_{r/2}(x_0) \cap \Omega = U$.



Now define

$$x_\epsilon = 2\epsilon\hat{e}_n$$

So the idea is after shrinking this region, we can shrink the ball centered at x_ϵ still lies within the larger ball

$$B_\epsilon(x_\epsilon) \subset B_r(x_0) \cap \Omega$$

Now define the perturbation on U:

$$u_\epsilon(x) = u(x_\epsilon)$$

Further, we can formally define its convolution:

$$v_\epsilon(x) := \underbrace{\phi_\epsilon}_{\text{smooth identity}} * u_\epsilon \in C^\infty(\overline{U})$$

Now we perform the calculation:

$$\lim_{\epsilon \downarrow 0} \|D^\alpha v_\epsilon - D^\alpha u\|_{L^p} \leq \lim_{\epsilon \downarrow 0} \|D^\alpha v_\epsilon - D^\alpha u_\epsilon\|_{L^p} + \lim_{\epsilon \downarrow 0} \|D^\alpha u_\epsilon - D^\alpha u\|_{L^p} = 0$$

Therefore, $v_\epsilon \to u$ in $W^{k,p}(\overline{U})$. Now, leveraging the boundedness of our boundary, we can cover the boundary with a finite net of open balls:

$$\partial\Omega = \bigcup_{i=1}^{N} B_{r_i/2}(x_i)$$

Because these balls need to overlap, we may define an interior closed set U_0 such that U_0 is a compact subset of Ω possessing the property that

$$\Omega = U_0 \cup \left(\bigcup_{i=1}^{N} \underbrace{B_{r_i/2}(x_i) \cap \Omega}_{U_i} \right) = \bigcup_{i=0}^{N} U_i$$

We know each set U_i is well behaved, and since $v_\epsilon \to u$ in $W^{k,p}(\overline{U})$, then there exists $v_j \in C^\infty(\overline{U_j})$ such that

$$\|v_j - u\|_{W^{k,p}(U_j)} < \delta$$

By the interior approximation, there exists $v_0 \in C^\infty(\overline{U_0})$ such that

$$\|v_0 - u\|_{W^{k,p}(U_j)} < \delta$$

Because Ω is covered by $\{U_i\}_{i=0}^N$, then by the interior approximation, we can let η_i be a partition of unit subordinate to this cover. Ultimately, we then we want to define an approximation over this subcover:

$$v = \sum_{i=0}^{N} v_i \eta_i \in C^\infty(\overline{\Omega})$$

Therefore, checking this has the desired properties:

$$\begin{aligned}
\|v - u\|_{W^{k,p}(\Omega)} &= \left\| \sum_{i=0}^{N} v_i \eta_i - u \right\|_{W^{k,p}(\Omega)} \\
&\leq \sum_{i=0}^{N} C \|v_i - u\|_{W^{k,p}(\Omega)} \\
&\leq CN\delta
\end{aligned}$$

■

19.4 Extensions and Traces

Lemma 19.15 (Existence of an Extension)**.** *Let $\Omega \subset\subset K \subset \mathbb{R}^n$ where K is compact and $\overline{\Omega} \subset K$. Let $\partial\Omega \in C^1$. Then there exists a bounded linear operator*

$$E : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$$

where $1 \leq p \leq \infty$ such that the following are true:

1. *$Eu = u$ a.e. $x \in \Omega$*
2. *$Eu = 0$ on K^c*
3. *$\|Eu\|_{W^{k,p}(\mathbb{R}^n)} \leq C_{p,\Omega,K} \|u\|_{W^{k,p}(\Omega)}$*

Proof. Since $C^\infty(\overline{\Omega})$ is dense in $W^{1,p}(\Omega)$, we may assume $u \in C^\infty(\overline{\Omega}) \cap W^{1,p}(\Omega)$. *Case:* Assume $\Omega = K = B^+ = \{x \in \mathbb{R}^b : \|x\| \leq 1, x_n \geq 0\}$. Define

$$E'u = \begin{cases} u(x_1, \dots, x_{n-1}, x_n) & x \in B^+ \\ u(x_1, \dots, x_{n-1}, -x_n) & x \in B^- \end{cases}$$

This is simply the mirror reflection over the $x_n = 0$ axis. This obviously extends continuously the function u. But if we want to make the derivative continuous, we define

$$Eu = \begin{cases} u(x_1, \dots, x_{n-1}, x_n) & x \in B^+ \\ -3u(x_1, \dots, x_{n-1}, -x_n) + 4u(x_1, \dots, x_{n-1}, \frac{-1}{2}x_n) & x \in B^- \end{cases}$$

Now denoting $u^+ = Eu|_{B^+}$ and $u^- = Eu|_{B^-}$ then we see still maintain the continuity condition

$$u^-|_{x_n=0} = u^+|_{x_n=0}$$

On the other hand, checking the derivative is also continuous:

$$\begin{aligned} \frac{\partial}{\partial x_n} u^-|_{x_n=0} &= 3\frac{\partial}{\partial x_n} u^-|_{x_n=0} - 2\frac{\partial}{\partial x_n} u^-|_{x_n=0} \\ &= \frac{\partial}{\partial x_n} u^-|_{x_n=0} \end{aligned}$$

Therefore the derivative is continuous! Checking the norm of our extension, we see:

$$\|Eu\|_{W^{1,p}(\mathbb{R}^n)} \leq C\,\|u\|_{W^{1,p}(B^+)}$$

Case: Assume a more interesting region $\bar{\Omega}$ compact with

$$\overline{\Omega} = \underbrace{\bigcup_{i=1}^{I} B_{r_i}(x_i)}_{\text{Interior Balls}} + \underbrace{\bigcup_{j=1}^{J} B_{r_j}(x_j)}_{\text{Exterior / Boundary Balls}}$$

such that $B_{r_i}(x_i) \subset \Omega$ and $x_j \in \partial\Omega$. By the assumptions of our extension, we know $\partial\Omega \in C^1$. Therefore, there must exists a diffeomorphism $\phi_j : B^+ \to K \in C^1$ such that $\phi_j^{-1} \in C^1$ and $\phi_j(B) = B_{\delta_j}(x_j)$. Lastly,

$$\phi_j(B^+) = B_{\delta_j}(x_j) \cap \Omega$$

Now let $\{\eta_i\}_{i=1}^I \cap \{\eta_j\}_{j=1}^J$ be a smooth partition of unit subordinate to the cover of $\overline{\Omega}$. Therefore, defining our extension:

$$\begin{aligned} u &= \sum_{i=1}^{I} \eta_i u + \sum_{j=1}^{J} \eta_j u \\ &= \underbrace{\sum_{i=1}^{I} \eta_i u}_{\text{Interior}} + \underbrace{\sum_{j=1}^{J} (\eta_j u) \circ \phi_j}_{\text{Use Extension Developed in First Case}} \end{aligned}$$

Therefore, the extension of this u becomes:

$$Eu = \sum_{i=1}^{I} \eta_i u + \sum_{j=1}^{J} E(\eta_j u) \circ \phi_j$$

Therefore, component-wise, we see:

$$\|Eu\| \leq C\,\|u\|_{W^{1,p}(\Omega)} + \sum_{j=1}^{J} \|u\|_{W^{1,p}(B_j)\cap\Omega}$$

■

Definition 19.16. *The* *boundary value of $u \in W^{1,p}(\Omega)$* *is defined as*

$$u|_{\partial\Omega} := \lim_{\epsilon \downarrow 0} Eu * \phi_\epsilon|_{\partial\Omega}$$

where ϕ_ϵ is a standard mollifier.

Theorem 19.17 (Trace)**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded, open subset with $\partial\Omega \in C^1$. Then there exists a bounded, linear operator T such that*

$$T : W^{1,p}(\Omega) \to L^p(\partial\Omega)$$

such that

1. *$u \in W^{1,p}(\Omega) \cap C^0(\overline{\Omega}), Tu = u|_{\partial\Omega}$*
2. *$\|Tu\|_{L^p(\partial\Omega)} \leq C_{p,\Omega}\,\|u\|_{W^{1,p}(\Omega)}$*

We refer to such an operator as a *Trace* *of u.*

Proof. Similar to the proof of the extension theorem, we will split the proof into two cases.
Let $\partial\Omega = \bigcup_{i=1}^{N} B_{r_i}(x_i)$ with the radii r_i small enough. The there exists functions ϕ_j that map each ball $B_{r_i}(x_i)$ diffeomorphically to the unit ball $B_1(0)$. Moreover

$$\phi_j(B_{r_i}(x_i)) = B_{r_i}(x_i)$$

$$\phi_j(B^+) = B_{r_i}(x_i) \cap \Omega$$

and for the boundary $\Gamma := \{x : \|x\| \leq 1 \text{ and } x_n = 0\}$

$$\phi_j(\Gamma) = \partial\Omega \cap B_{r_i}(x_i)$$

Therefore, we only need to perform calculations on each of these unit balls $B_1(0)$ in order to reach our conclusion!

In order to prove (2), that is, for any $u \in C_c^\infty(B^+)$,

$$\|Tu\|_{L^p(\Gamma)} \leq C\,\|u\|_{W^{1,p}(B^+)}$$

This is equivalent to proving

$$\int_{\mathbb{R}^{n-1}} |u(x_1, \ldots, x_{n-1}, 0)|^p dx_1 \wedge \ldots \wedge x_{n-1} \leq C\,\|u\|^p_{W^{1,p}(B^+)}$$

Now observe,

$$\begin{aligned}
\int_{\mathbb{R}^{n-1}} |u(x_1, \ldots, x_{n-1}, 0)|^p dx_1 \wedge \ldots \wedge dx_{n-1} &= \int_{\mathbb{R}^{n-1}} \left| \int_0^1 \frac{d}{dx_n} |u(x_1, \ldots, x_{n-1}, x_n)|^p dx_n \right| (dx_1 \wedge \ldots \wedge dx_{n-1}) \\
&= \int_{\mathbb{R}^n} p|u(x)|^{p-1} |\nabla u| dx \\
&\leq p \int_{\mathbb{R}^n} \left(\frac{|u(x)|^p}{q} + \frac{|\nabla u|^p}{p} \right) dx && \text{Young} \\
&= C_p \int_{B^+} |u|^p + |\nabla u|^p dx \\
&\leq C_p\,\|u\|^p_{W^{1,p}(B^+)}
\end{aligned}$$

This takes care of the model case! In order to extend to the more general region Ω, we let $\{\eta_j\}_j$ be a smooth partition of unity

$$\Omega \subset B_{r_j}(x_j)$$

Therefore, we can decompose $u \in C^\infty(\overline{\Omega})$ as:

$$u = \sum_{j=1}^{N} \eta_j u$$

Clearly,

$$\begin{aligned}
\|u\|_{L^p(\partial\Omega)} &\le \sum_{j=1}^{N} \|\eta_j u\|_{L^p(\partial\Omega \cap B_{r_j}(x_j))} \\
&= \sum_{j=1}^{N} \|\eta_j u \circ \phi_j\|_{L^p(\Gamma)} && \text{Map to model case} \\
&\le C_p \sum_{j=1}^{N} \|\eta_j u \circ \phi_j\|_{W^{1,p}(B^+)} \\
&\le C_p \sum_{j=1}^{N} \|\eta_j u\|_{W^{1,p}(B_{r_j}(x_j) \cap \Omega)} \\
&\le C_p \|u\|_{W^{1,p}(\Omega)}
\end{aligned}$$

By density, this property can be extended to any function that is smooth up to the boundary!

Secondly, we need to prove that this operator can be defined to be equal to u on the boundary. Let $u_m \in C^\infty(\overline{\Omega})$ such that

$$u_m \to u \text{ in } W^{1,p}(\Omega)$$

By the previous claim, the L^p norm on the boundary controlled by the interior implies

$$\limsup_{m,n\to\infty} \|u_m - u_n\|_{L^p(\partial\Omega)} \le C_p \lim_{m,n\to\infty} \|u_m - u_n\|_{W^{1,p}(\Omega)} = 0$$

Therefore, $\{u_m\}$ is a Cauchy sequence in $L^p(\partial\Omega)$. Now, define

$$u = \lim_{m\to\infty} u_m|_{\partial\Omega}$$

Now we need to verify that for $u \in C^0(\overline{\Omega}) \cap W^{1,p}(\Omega)$,

$$Tu = u|_{\partial\Omega}$$

The canonical approach is to choose

$$Tu = \lim_{\epsilon\downarrow 0} Eu * \phi_\epsilon$$

To show that this choose in fact equals u on the boundary, we recall that this convolution $Eu * \phi_\epsilon$ converges to u uniformly on a compact set $\partial\Omega$. Since the limit $Eu * \phi_\epsilon$ is unique and continuous, then we can conclude

$$\lim_{\epsilon\downarrow 0} Eu * \phi_\epsilon = u|_{\partial\Omega}$$

■

Not gonna lie, these lectures are getting harder to sit through. Pacing is getting faster and explanations are getting more and more hand-wavy. I'm having to pause in order to incorporate spoken details into the notes because otherwise the board notes are reaching incomprehensible.

19.5 Compactly Supported Sobolev Spaces

Definition 19.18. *Given $\Omega \subset \mathbb{R}^n$ a bounded, open set. We define*

$$W_0^{1,p}(\Omega) = \{u \in W^{1,p}(\Omega) : u|_{\partial\Omega} = 0\} = \overline{C_c^\infty(\Omega)}$$

where closure is taken with respect to the norm $\|\cdot\|_{W^{1,p}(\Omega)}$.

Question 19.19. *Why is this space useful?*

Example 19.20. *Suppose $u, v \in W^{1,p}(\Omega)$. Then we have the equivalent statements:*

$$u|_{\partial\Omega} = v|_{\partial\Omega} \iff u - v \in W_0^{1,p}(\Omega)$$

Remark 19.21. *In general, $W_0^{1,p}(\Omega) \subsetneq W^{1,p}(\Omega)$. An interesting class of spaces arrises when this is no longer true.*

Theorem 19.22. $1 \leq p < \infty$,

$$W^{1,p}(\mathbb{R}^n) = W_0^{1,p}(\mathbb{R}^n)$$

Proof. Let $u \in W^{1,p}(\mathbb{R}^n)$ and define the bump function

$$\eta_k(x) = \begin{cases} 1 & |x| \leq k \\ \text{smooth sigmoid-like} & k \leq |x| \leq k+1 \\ 0 & |x| \geq k \end{cases}$$

If we consider $u\eta_k \in C_c^\infty(\mathbb{R}^n) \cap W^{1,p}(\mathbb{R}^n)$ with $|D\eta_k| \leq C$ for all k. Observe:

$$\begin{aligned}
\|u - u\eta_k\|_{W^{1,p}(\mathbb{R}^n)} &= \left[\int |u|^p \underbrace{|1-\eta_k|^p}_{\text{bounded by } 1}\right]^{\frac{1}{p}} + \left[\int |Du|^p \underbrace{|1-\eta_k|^p}_{\leq 1}\right]^{\frac{1}{p}} + \left[\int |u|^p \underbrace{|D\eta_k|^p}_{\leq C}\right]^{\frac{1}{p}} \\
&\leq \left[\lim_{k\to\infty} \int |u|^p |1-\eta_k|^p\right]^{\frac{1}{p}} + \left[\lim_{k\to\infty} \int |Du|^p |1-\eta_k|^p\right]^{\frac{1}{p}} + \left[\lim_{k\to\infty} \int |u|^p |D\eta_k|^p\right]^{\frac{1}{p}} \quad \text{Dominated Converg} \\
&\leq \lim_{k\to\infty} \int |u|^p \mathcal{X}_{k\leq|x|\leq k+1} + \lim_{k\to\infty} \int |Du|^p \mathcal{X}_{k\leq|x|\leq k+1} + C \lim_{k\to\infty} \int |u|^p \mathcal{X}_{k\leq|x|\leq k+1} \\
&= 0
\end{aligned}$$

since u, Du are integrable and therefore finite almost everywhere. ■

19.6 Sobolev's Inequality

In order to proceed, let's recall some useful facts/ generalizations:

Lemma 19.23 (Generalized Hölder Inequality)**.** *Given $u_i \in L^{p_i}(\mathbb{R}^n)$ with $1 \leq p_i \leq \infty$ and $1 \leq i \leq k$. Then if*

$$\frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_k} = 1$$

then

$$\int \prod_{i=1}^{k} |u_i| \leq \prod_{i=1}^{k} \|u_i\|_{L^{p_i}(\mathbb{R}^n)}$$

Lemma 19.24. *Given $n \geq 2$, $u_i \in L^{n-1}(\mathbb{R}^{n-1})$ with the deleted element vector denoted*

$$\hat{x}_i = (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in \mathbb{R}^{n-1}$$

Suppose we define:

$$u(x) := \prod_{i=1}^{n} u_i(\hat{x}_i)$$

Then

$$\|u\|_{L^1(\mathbb{R}^n)} \leq \prod_{i=1}^{n} \|u_i\|_{L^{n-1}(\mathbb{R}^{n-1})}$$

Proof. Basis: $n = 2$. Then

$$u(x) = u_1(x_2) \cdot u_2(x_1)$$

Observe:

$$\int |u| dx_1 dx_2 = \int |u_1(x_2)| dx_2 \int |u_2(x_1)| dx_1$$

$$\iff \|u\|_{L^1} = \|u_1\|_{L^1} \|u_2\|_{L^1}$$

Inductive Step: Fix x_{n+1} and let

$$u(x) = \prod_{i=1}^{n} u_i(\hat{x}_i) \cdot u_{n+1}(\hat{x}_n)$$

Applying Hölder's Inequality:

$$\int |u(x)| dx \leq \|u_{n+1}\|_{L^{n-1}(\mathbb{R}^n)} \left[\int \prod_{i=1}^{n} \underbrace{|u_i(\hat{x}_i)|^{\frac{n}{n-1}}}_{v_i} d\hat{x}_i \right]^{\frac{n-1}{n}}$$

Notice $\hat{x}_i = (\hat{x}_i', x_{n+1}) \in \mathbb{R}^n$ and therefore, the vector $\hat{x}_i' \in \mathbb{R}^{n-1}$, and therefore a suitable construction to use the inductive hypothesis over. Observe,

$$v_i(\hat{x}_i') = |u_i(\hat{x}_i', x_{n+1})|^{\frac{n}{n-1}}$$

Therefore, letting $V = \prod_{i=1}^{n} v_i$ such that $V : \mathbb{R}^n \to \mathbb{R}$, we can apply the inductive hypothesis:

$$\|V\|_{L^1(\mathbb{R}^n)} \leq \prod_{i=1}^{n} \|v_i\|_{L^{n-1}(\mathbb{R}^n)}$$

We can return to our previous calculation to see:

$$\begin{aligned}
\int_{\mathbb{R}^n} u(x', x_n) dx' &\leq \|u_{n+1}\|_{L^n(\mathbb{R}^n)} \left[\int \prod_{i=1}^{n} |u_i(\hat{x}_i)|^{\frac{n}{n-1}} d\hat{x}_i \right]^{\frac{n-1}{n}} \\
&\leq \|u_{n+1}\|_{L^n(\mathbb{R}^n)} \left[\prod_{i=1}^{n} \int |u_i(\hat{x}_i')|^n d\hat{x}_i' \right]^{\frac{1}{n}}
\end{aligned}$$

Integrating both side with respect to x_{n+1}, we use the generalized Hölder Inequality with $(\frac{1}{n} + \ldots + \frac{1}{n}) = 1$:

$$\|u\|_{L^1(\mathbb{R}^n)} = \int_{\mathbb{R}^{n+1}} |u(x)| dx \leq \|u_{n+1}\|_{L^n(\mathbb{R}^n)} \prod_{i=1}^{n} \|u_i\|_{L^n(\mathbb{R}^n)}$$

Therefore, by the principle of mathematical induction, we have shown that the case of n implies the case with $n + 1$, concluding the proof. ∎

Now, we want to develop an inequality of the form:

$$\|u\|_{L^q(\mathbb{R}^n)} \leq C\,\|Du\|_{L^p(\mathbb{R}^n)}$$

where $u \in C_c^\infty(\mathbb{R}^n)$
To reason what the relationship between q and p is, we can leverage a rescaling argument such that $u(\lambda x)$, in which one concludes

$$q = p* = \frac{np}{n-p}$$

Theorem 19.25 (Sobolev's Inequalities). *Suppose $u \in W^{1,p}(\mathbb{R}^n)$, then $u \in L^{p*}(\mathbb{R}^n)$, and*

$$\|u\|_{L^{p*}(\mathbb{R}^n)} \leq C\,\|Du\|_{L^p(\mathbb{R}^n)}$$

Proof. Case: Let $p = 1$. Then

$$\begin{aligned} |u(x)| &\leq \int_{-\infty}^{x_i} \left|\frac{\partial u}{\partial x_i}\right| dx_i \\ &\leq \underbrace{\int |Du(x_1, \ldots, x_2, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_n)| dy_i}_{W(\hat{x}_i)} \end{aligned}$$

$$\implies |u(x)|^{\frac{n}{n-1}} \leq \prod_{i=1}^n W(\hat{x}_i)^{\frac{1}{n-1}}$$

Integrating with respect to x:

$$\begin{aligned} \int |u(x)|^{\frac{n}{n-1}} dx &\leq \left\| \prod_{i=1}^n W(\hat{x}_i)^{\frac{1}{n-1}} \right\|_{L^1(\mathbb{R}^n)} \\ &\leq \prod_{i=1}^n \left\| W(\hat{x}_i)^{\frac{1}{n-1}} \right\|_{L^{n-1}(\mathbb{R}^{n-1})} \\ &= \prod_{i=1}^n \left[\int_{\mathbb{R}^{n-1}} W(\hat{x}_i) d\hat{x}_i \right]^{\frac{1}{n-1}} \\ &= \|Du\|_{L^1(\mathbb{R}^n)}^{\frac{n}{n-1}} \end{aligned}$$

Therefore,

$$\|u\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \leq \|Du\|_{L^1(\mathbb{R}^n)}$$

Case: $1 < p < n$. Let $v = |u|^\gamma \in C^1$ where γ is yet to be determined, but assumed to be greater than 1. Then

$$\int v^{\frac{n}{n-1}} \leq \int |Dv| dx$$

$$\begin{aligned} \iff \int |u|^{\frac{\gamma n}{n-1}} &\leq \int \gamma |u|^{\gamma-1} |Du| dx \\ &\leq \gamma \left[\int |u|^{\frac{p(\gamma-1)}{p-1}} \right]^{\frac{p-1}{p}} \left[\int |Du|^p dx \right]^{\frac{1}{p}} \end{aligned}$$

Choosing $\frac{\gamma n}{n-1} = \frac{p(\gamma-1)}{p-1} \implies \gamma = \frac{p(n-1)}{n-p}$, we see:

$$\|u\|_{L^{p*}} \leq \gamma\,\|Du\|_{L^p}$$

∎

Corollary 19.25.1. *$1 \leq p < \infty, p < n$ and $\Omega \subset \mathbb{R}^n$ is a bounded, open subset with $\partial\Omega \in C^1$. Then*

$$\|u\|_{L^{p*}(\Omega)} \leq C_{p,n,\Omega} \|u\|_{W^{1,p}(\Omega)}$$

Proof. Let $\bar{u} = Eu \in W^{1,p}(\mathbb{R}^n)$ be an extension of u such that

$$\|\bar{u}\|_{W^{1,p}(\mathbb{R}^n)} \leq C_{\Omega,n,p} \|u\|_{W^{1,p}(\Omega)}$$

By the Sobolev Inequality,

$$\|\bar{u}\|_{L^{p*}(\mathbb{R}^n)} \leq \|D\bar{u}\|_{L^p(\mathbb{R}^n)}$$

More obviously, $\|u\|_{L^{p*}(\mathbb{R}^n)} \leq \|\bar{u}\|_{L^{p*}(\mathbb{R}^n)}$. Therefore, we arrive at the chain:

$$\|u\|_{L^{p*}(\mathbb{R}^n)} \leq \|\bar{u}\|_{L^{p*}(\mathbb{R}^n)} \leq \|D\bar{u}\|_{L^p(\mathbb{R}^n)} \leq \|\bar{u}\|_{W^{1,p}(\mathbb{R}^b)} \leq C_{\Omega,n,p} \|u\|_{W^{1,p}(\Omega)}$$

■

19.7 Applications of the Sobolev Inequality

Example 19.26. *Let $p = 1$ with $\Omega \subset \mathbb{R}^n$ bounded open set with C^1 boundary. Define*

$$u_\epsilon = \begin{cases} 1 & x \in \Omega \\ 1 - \frac{dist(x,\Omega)}{\epsilon} & 0 < dist(x,\Omega) < \epsilon \\ 0 & dist(x,\Omega) > \epsilon \end{cases}$$

Then u_ϵ. Observe, we end up with the chain

$$\left[\int \mathcal{X}_\Omega^{\frac{n}{n-1}}\right]^{\frac{n-1}{n}} \leq \|u_\epsilon\|_{L^{\frac{n}{n-1}}} \leq \|Du\|_{L^1} \leq c\int \frac{1}{\epsilon}dx = c\underbrace{\frac{|\{x : 0 < dist(x,\Omega) < \epsilon\}|}{\epsilon}}_{H^{n-1}(\partial\Omega)} \approx cperimeter$$

This is known as the Isoperimetric Inequality

Question 19.27. *Recall as a consequence of the Sobolev Inequality, When $p < n$, $W_{1,p}(\mathbb{R}^n) \subset L^{p*}(\mathbb{R}^n)$ where $p* = \frac{np}{n-p}$. This begs the question: What happens when $p > n$?*

Morrey's Inequality says when $p > n$, $W^{1,p} \subset C^{0,\alpha}(\mathbb{R}^n)$ for some $0 < \alpha < 1$. What is this α?

Definition 19.28. *A function $f : \Omega \to \mathbb{R}$ is Hölder α-continuous provided there exists a constant C such that*

$$|f(x) = f(y)| \leq C|x - y|^\alpha$$

for all $x, y \in \Omega$.

Definition 19.29. *The space $C^{0,\alpha}(\Omega) = C^\alpha(\Omega) =$ is the space of Hölder α-continuous functions equipped with the finite norm*

$$\|f\|_{C^\alpha(\Omega)} = \sup_{x\in\Omega} |f(x)| + \sup_{x,y\in\Omega, x\neq y} \frac{|f(x) - f(y)|}{|x - y|}$$

Theorem 19.30 (Morrey's Inequality)**.** *Assume $f \in W^{1,p}(\mathbb{R}^n), p > n$. Then*

$$|f(x) - f(y)| \leq \left(\frac{2np}{p-n}\right) |x - y|^{1-\frac{n}{p}} \|Du\|_{L^p}$$

Before we prove this inequality, we need the following lemma:

Lemma 19.31 (Morrey)**.** *Assume $f \in C^\infty(\mathbb{R}^n) \cap W^{1,p}(\mathbb{R}^n)$ with $p > n$. Define*

$$Q_\ell(x) := \text{ cube centered at } x \text{ with side length } \ell$$

And define the function

$$f_{Q_\ell(x)} = \frac{1}{|Q_\ell(x)|} \int_{Q_\ell(x)} f(y) dy$$

If $y \in Q_\ell(x)$, then

$$|f(y) - f_{Q_\ell(x)}| = \left| \frac{1}{|Q_\ell(x)|} \int_{Q_\ell(x)} f(y) dy \right| \leq \frac{np}{p-n} \ell^{1-\frac{n}{p}} \|Df\|_{L^p(Q_\ell(x))}$$

Proof. (Of Morrey's Inequality)
Let $z = \frac{x+y}{2}$ and $\ell = |x - y|$. Then it follows

$$x, y \in Q_\ell(z)$$

We can therefore apply Morry's lemma to both x, y to get:

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_{Q_\ell(z)}| + |f(y) - f_{Q_\ell(z)}| \\ &\leq \frac{2np}{p-n} |x - y|^{1-\frac{n}{p}} \|Du\|_{L^p(\mathbb{R}^n)} \end{aligned}$$

■

Now we return to prove Morrey's Lemma:

Proof. (Morrey's Lemma)
Writing our lemma's conclusion in a slightly different way:

$$\frac{np}{p-n} \ell^{1-\frac{n}{p}} \|Df\|_{L^p(Q_\ell(x))} = \frac{np}{p-n} \ell \left(\frac{1}{\ell^n} \int |Df|^p \right)^{\frac{1}{p}} = \frac{np}{p-n} \ell \left(\frac{1}{|Q_\ell(x)|} \int_{Q_\ell(x)} |Du|^p \right)^{\frac{1}{p}}$$

Therefore, we see that latter portion is more approachable in calculation. Specifically, we need to demonstrate the inequality,

$$\left| \frac{1}{|Q_\ell(x)|} \int_{Q_\ell(x)} f(y) dy \right| \leq \frac{np}{p-n} \ell \left(\frac{1}{|Q_\ell(x)|} \int_{Q_\ell(x)} |Du|^p \right)^{\frac{1}{p}}$$

Observe, by the fundamental theorem of Calculus ,

$$\begin{aligned} f(z) - f(y) &= \int_0^1 \frac{d}{d\theta} f(\theta z + (1-\theta) y) \ d\theta \\ &= \int_0^1 Df(\theta z + (1-\theta) y) * (z - y) d\theta \end{aligned}$$

Therefore,

$$\begin{aligned}
|f(y) - f_{Q_\ell(x)}| &= \left| \frac{1}{|Q_\ell(x)|} \int_{Q_\ell(x)} (f(z) - f(y))\ dz \right| \\
&= \left| \left| \frac{1}{|Q_\ell(x)|} \int_{Q_\ell(x)} \int_0^1 Df(\theta z + (1-\theta)y) * (z-y) d\theta \right| dz \right| \\
&\le \sum_{i=1}^n \int_0^1 \frac{1}{\ell^n} \int_{Q_\ell(x)} \left| \frac{\partial u}{\partial x_i}(\theta z + (1-\theta)y) \right| |z_i - y_i|\ dz\ d\theta \\
&= \sum_{i=1}^n \frac{1}{\ell^{n-1}} \int_0^1 \int_{Q_{\ell\theta}(\theta x + (1-\theta)y} \left| \frac{\partial u}{\partial x_i}(W) \right| \frac{dW}{\theta^n} d\theta \\
&\le \sum_{i=1}^n \frac{1}{\ell^{n-1}} \int_0^1 \left[\int_{Q_{\ell\theta}(\theta x + (1-\theta)y} \left| \frac{\partial u}{\partial x_i} \right|^p \right]^{\frac{1}{p}} \left[(\ell\theta)^{\frac{n}{q}} \right] \frac{d\theta}{\theta^n} && \text{Hölder} \\
&\le \|Du\|_{L^p(Q_\ell(x))} \ell^{1-\frac{n}{p}} \int_0^1 \theta^{-\frac{n}{p}} d\theta \\
&= \|Du\|_{L^p(Q_\ell(x))} \ell^{1-\frac{n}{p}} \frac{np}{p-n}
\end{aligned}$$

■

Corollary 19.31.1. *Let $p > n$ and $f \in W^{1,p}(\mathbb{R}^n)$. Then*

$$\sup_{x \in \mathbb{R}^n} |f(x)| \le C_{n,p} \|f\|_{W^{1,p}(\mathbb{R}^n)}$$

Proof. Given an x, we can look for a cube $Q_1(z)$ that contains this x. Now

$$\begin{aligned}
|f(x)| &\le |f(x) - f_{Q_1(z)}| + |f_{Q_1(z)}| \\
&\le \frac{np}{p-n} \|Df\|_{L^p(Q_1(z))} + \int_{Q_1(z)} |f| dx && \text{Morrey's Lemma} \\
&\le \frac{np}{p-n} \|Df\|_{L^p(Q_1(z))} + \|f\|_{L^p(Q_1(z))} \\
&\le \max\left\{ \frac{np}{p-n}, 1 \right\} \|f\|_{W^{1,p}(\mathbb{R}^n)}
\end{aligned}$$

■

This tells you that when $p > n$, $W^{1,p}(\mathbb{R}^n) \subset C^\alpha(\mathbb{R}^n)$ for

$$\alpha = 1 - \frac{n}{p}$$

Definition 19.32. *$f : \mathbb{R}^n \to \mathbb{R}$ is <u>differentiable</u> at x provided there exists a linear map $\ell : \mathbb{R}^n \to \mathbb{R}$ such that*

$$\lim_{y \to x} \frac{|f(y) - f(x) - L(x-y)|}{|x-y|} = 0$$

If such an L exists, L is called the derivative of f at x.

Theorem 19.33. *If $f \in W^{1,p}_{loc}(\mathbb{R}^n)$, then f is differentiable almost everywhere and it's derivative is equal to the weak derivative.*

Proof. By the Lebesgue Differentiation Theorem, for a.e. $x \in \mathbb{R}^n$

$$\lim_{\ell \downarrow 0} \frac{1}{|Q_\ell(x)|} \int_{Q_\ell(x)} |Df(z) - Df(x)|^p dz = 0$$

Fix x satisfying the above property. Then define the function F by

$$F(y) := f(y) - f(x) - Df(x)(y-x)$$

Then $F \in W^{1,p}_{loc}(\mathbb{R}^n)$ with $p > n$. Therefore, we can apply Morrey's Lemma to get:

$$|F(y) - F(x)| \leq C\ell \left(\frac{1}{|Q_\ell(x)|} \int_{Q_\ell(x)} |DF(z)|^p dz \right)^{\frac{1}{p}}$$

Letting $\ell = 2|x-y| \implies y \in Q_\ell(x)$. Therefore, since $F(x) = 0$, then we can write:

$$\begin{aligned}|f(y) - f(x) - Df(x)(y-x)| &\leq C\ell \left(\frac{1}{|Q_\ell(x)|} \int_{Q_\ell(x)} |DF(z)|^p dz \right)^{\frac{1}{p}} \\ &\leq 2C|x-y| \left[\frac{1}{|Q_\ell(x)|} \int_{|Q_\ell(x)|} |Df(z) - Df(x)|^p dz \right]^{\frac{1}{p}} \to 0 \text{ as } \ell \to 0\end{aligned}$$

Therefore, we can rearrange the chain above to achieve

$$\lim_{\ell \downarrow 0} \frac{|f(y) - f(x) - Df(x)(y-x)|}{|x-y|} = 0$$

Moreover, the weak derivative simultaneously serves as the derivative L. ∎

Question 19.34. *So the Sobolev inequality answers what happens when $p < n$ and Morrey's answers for the case $p > n$. Now what happens in the borderline case in which $p = n$?*

Lemma 19.35 (Poincaré)**.** *Let $\Omega \subset \mathbb{R}^n$ be bounded with $1 \leq p < \infty$. If $f \in W^{1,p}_0(\Omega)$, then*

$$\int_\Omega |f|^p \ dx \leq C_p (diam\ \Omega)^p \int_\Omega |Df|^p \ dx$$

where $diam\ \Omega := \sup_{x,y \in \Omega} \|x - y\|$. Further:

$$\left(\frac{1}{|\Omega|} \int_\Omega |f|^p \ dx \right)^{\frac{1}{p}} \leq C_p (diam\ \Omega)^p \left(\frac{1}{|\Omega|} \int_\Omega |Df|^p \ dx \right)^{\frac{1}{p}}$$

Remark 19.36. *Notice it's necessary for f to be zero on the boundary. Otherwise, you arrive at a contradiction depending on the region chosen.*

Proof. Without loss of generality, we nod in the direction of a density argument, by which we may assume $f \in C^\infty_c(\Omega)$. Take any $y \in \Omega$ and write $y = (y_1, y_2, \ldots, y_n)$. Then

$$\Omega \subset \prod_{i=1}^n [y_i - diam\ \Omega_i, y_i + diam\ \Omega_i] := \prod_{i=1}^n [a_i, b_i]$$

Since f has zero boundary, we can write:

$$f(x) = \int_{a_i}^{x_i} \frac{\partial}{\partial x_i} f(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n) dt$$

This implies

$$|f(x)| \leq \int_{a_i}^{b_i} |Df(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)| dt$$

By Hölder's inequality, we see:

$$\int_{a_i}^{b_i} |Df(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)| dt \leq (b_i - a_i)^{1-\frac{1}{p}} \left[\int_{a_i}^{b_i} |Df(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)|^p dt \right]^{\frac{1}{p}}$$

Since $b_i - a_i \leq 2diam\ \Omega_i$, we see:

$$(b_i - a_i)^{1-\frac{1}{p}} \left[\int_{a_i}^{b_i} |Df(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)|^p dt \right]^{\frac{1}{p}} \leq (2diam\ \Omega)^{1-\frac{1}{p}} \left[\int_{a_i}^{b_i} |Df(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)|^p dt \right]^{\frac{1}{p}}$$

That is,

$$|f(x)| \leq (2diam\ \Omega)^{1-\frac{1}{p}} \left[\int_{a_i}^{b_i} |Df(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)|^p dt \right]^{\frac{1}{p}}$$

Case: When $i = 1$, then we see:

$$\begin{aligned}
\int_\Omega |f|^p\ dx &\leq \int_\Omega (2diam\ \Omega)^{p-1} \int_{a_i}^{b_i} |Df(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)|^p dt\ dx \\
&= \int_\Omega (2diam\ \Omega)^{p-1} \int_{a_1}^{b_1} |Df(t, x_2, \dots, x_n)|^p dx_1\ dt\ dx_2\ \dots dx_n && \text{Fubini} \\
&= \int_\Omega (2diam\ \Omega)^p |Df(x)|^p\ dx \\
&= 2^p (diam\ \Omega)^p \int_\Omega |Df|^p\ dx
\end{aligned}$$

■

Question 19.37. *How can we extend this idea onto functions that don't have a zero boundary?*

Lemma 19.38 (Poincaré's Inequality on Cubes)**.** *Let $\Omega \subset \mathbb{R}^n$ be an open subset. Suppose $f \in W^{1,p}_{loc}(\Omega)$ with $1 \leq p < \infty$. Then*

$$\left(\frac{1}{|Q_\ell(x)|} \int_{Q_\ell(x)} |f(x) - f_{Q_\ell(x)}|^p\ dx \right)^{\frac{1}{p}} \leq C_{n,p} \ell \left[\frac{1}{|Q_\ell(x)|} \int_{Q_\ell(x)} |Df(y)|^p\ dy| \right]^{\frac{1}{p}}$$

for any $Q_\ell(x) \subset\subset \Omega$. Multiplying both sides by the volume, we see this statement is equivalent to:

$$\left(\int_{Q_\ell(x)} |f(x) - f_{Q_\ell(x)}|^p\ dx \right)^{\frac{1}{p}} \leq C_{n,p} \ell \left[\int_{Q_\ell(x)} |Df(y)|^p\ dy| \right]^{\frac{1}{p}}$$

Proof. Without loss of generality, we may assume $f \in C^\infty(\Omega)$. Take $z, y \in Q_\ell(x) = \prod_{i=1}^n [a_i, b_i]$.

$$\begin{aligned}
|f(x) - f(y)| &= |f(z_1, z_2, \dots, z_n) - f(z_1, z_2, \dots, y_n)| \\
&\quad + |f(z_1, z_2, \dots, z_{n-1}, y_n) - f(z_1, z_2, \dots, y_{n-1}, y_n)| \\
&\quad + \dots + |f(z_1, y_2, \dots, y_n) - f(y_1, y_2, \dots, y_n)| \\
&\leq \sum_{i=1}^n \int_{a_i}^{b_i} |Df(z_1, z_2, \dots, z_{i-1}, t, y_{i+1}, \dots, y_n)|\ dt \\
&\leq \sum_{i=1}^n (b_i - a_i)^{1-\frac{1}{p}} \left(\int_{a_i}^{b_i} |Df(z_1, z_2, \dots, z_{i-1}, t, y_{i+1}, \dots, y_n)|^p\ dt \right)^{\frac{1}{p}} && \text{Hölder}
\end{aligned}$$

Next, we make use of the following lemma:

Lemma 19.39.
$$\left(\sum_{i=1}^{N} A_i\right)^p \le n^p \left(\sum_{i=1}^{N} A_i^p\right)$$
provided $A_i \ge 0$.

Therefore, raising both sides to the pth power:

$$\begin{aligned}
|f(x)-f(y)|^p &\le \left[\sum_{i=1}^{n} (b_i-a_i)^{1-\frac{1}{p}} \left(\int_{a_i}^{b_i} |Df(z_1,z_2,\dots,z_{i-1},t,y_{i+1},\dots,y_n)|^p \, dt\right)^{\frac{1}{p}}\right]^{\frac{1}{p}} \\
&\le n^p \sum_{i=1}^{n} \underbrace{(b_i-a_i)^{p-1}}_{\ell} \int_{a_i}^{b_i} |Df(z_1,z_2,\dots,z_{i-1},t,y_{i+1},\dots,y_n)|^p \, dx \qquad \text{By Lemma}
\end{aligned}$$

Therefore,
$$|f(x)-f(y)|^p \le n^p \ell^{p-1} \sum_{i=1}^{n} \int_{a_i}^{b_i} |Df(z_1,z_2,\dots,z_{i-1},t,y_{i+1},\dots,y_n)|^p \, dx$$

We can now integrate both sides over the region $Q_\ell(x)$:

$$\begin{aligned}
\int_{Q_\ell(x)} |f(z)-f_{Q_\ell(x)}|^p \, dx &= \int_{Q_\ell(x)} \left| ⨍_{Q_\ell(x)} f(z)-f(y)dy \right|^p \, dz \\
&\le \int_{Q_\ell(x)} ⨍_{Q_\ell(x)} |f(z)-f(y)|^p \, dy \, dz \qquad \int_Q \underbrace{|F|}_{p} \underbrace{dx}_{q} \le \int_Q \\
&= ⨍_{Q_\ell(x)} \int_{Q_\ell(x)} |f(z)-f(y)|^p \, dz \, dy \\
&= \frac{n^p \ell^{p-1}}{|Q_\ell(x)|} \sum_{i=1}^{n} \int_{Q_\ell(x)} \int_{Q_\ell(x)} \int_{a_i}^{b_i} |Df(z_1,\dots,z_{i-1},t,y_{i+1},\dots,y_n)|^p \, dt \, dz \, dy \\
&= \frac{n^p \ell^{p-1}}{|Q_\ell(x)|} \sum_{i=1}^{n} (b_i-a_i)^{\ell} \int_{Q_\ell(x)} \int_{Q_\ell(x)} |Df(u)|^p \, du \, dw \\
&= n^p \ell^p \sum_{i=1}^{n} \int_{Q_\ell(x)} |Df(u)|^p \, du \\
&= n^{p+1} \ell^p \int_{Q_\ell(x)} |Df(u)|^p \, du
\end{aligned}$$

■

We can push this more general domains, but they require more care when dealing with the boundary of these functions.

Question 19.40. *Can we identify a class of functions for which $W^{1,n}(\mathbb{R}^n)$ is embedded?*

Proof. If $f \in W^{1,n}(\mathbb{R}^n)$, then for every cube Q

$$\begin{aligned}
\int_Q |f - f_Q| dx &\leq \left(\fint_Q |f - f_Q|^n \right)^{\frac{1}{n}} && \text{Hölder} \\
&\leq C\ell \left(\frac{1}{\ell^n} \int_Q |Df|^n \right)^{\frac{1}{n}} && \text{Poincaré} \\
&= C \left(\fint_Q |Df|^n \right)^{\frac{1}{n}} < \infty
\end{aligned}$$

Therefore, defining a norm:

$$\|f\|_{BMO} = \sup_{Q \subset \mathbb{R}^n} \int_Q |f - f_Q| dx$$

We are forced to see $W^{1,n}(\mathbb{R}^n) \subset BMO(\mathbb{R}^b)$ where the space BMO stands for Bounded Mean Oscillation. ■

20 Fourier Transforms

Definition 20.1. *Suppose $f \in L^1(\mathbb{R}^n)$. we define the Fourier Transform by*

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x)e^{-2\pi i\xi x}\ dx$$

20.1 Properties of the Fourier Transform

Lemma 20.2. *For any $f \in L^1(\mathbb{R}^n)$, $\hat{f}$ is continuous.*

Proof. Let $\xi_j \to \xi$. Then

$$\begin{aligned}
\lim_{\xi_j \to \xi} \hat{f}(\xi_j) &= \lim_{\xi_j \to \xi} \int_{\mathbb{R}^n} f(x)e^{-2\pi\xi_j x}\ dx \\
&= \int_{\mathbb{R}^n} f(x) \lim_{\xi_j \to \xi} e^{-2\pi\xi_j x}\ dx && \text{Dominatd Convergence} \\
&= \int_{\mathbb{R}^n} f(x)e^{-2\pi\xi x}\ dx
\end{aligned}$$

■

Lemma 20.3. $f \in L^1 \implies \hat{f} \in L^\infty$.

Proof.

$$\left\|\hat{f}\right\|_{L^\infty(\mathbb{R}^n)} = \int_{\mathbb{R}^n} |f(x)e^{-2\pi i\xi x}|\ dx \leq \int_{\mathbb{R}^n} |f|dx < \infty$$

■

Lemma 20.4. $f \to \hat{f}$ *is linear.*

Definition 20.5. $\tau_h f(x) := f(x+h)$

Lemma 20.6. $\widehat{\tau_h f} = e^{2\pi i\xi h}\widehat{f}(\xi)$

Proof.

$$\begin{aligned}
\widehat{\tau_h f} &= \int_{\mathbb{R}^n} (\tau_h f)(x)e^{-2\pi i\xi x}\ dx \\
&= \int_{\mathbb{R}^n} f(y)e^{-2\pi i\xi(y-h)}\ dy \\
&= e^{2\pi i\xi h}\int_{\mathbb{R}^n} f(y)e^{-2\pi i\xi y}\ dy \\
&= e^{2\pi i\xi h}\widehat{f}(\xi)
\end{aligned}$$

■

Remark 20.7. *Clearly, $\hat{f}(0) = \int_{\mathbb{R}^n} f(x)\ dx$*

Definition 20.8. $\delta_\lambda f(x) := f(\lambda x)$

Lemma 20.9.

$$\widehat{\delta_\lambda f}(\xi) = \frac{1}{\lambda^n}(\delta_{\frac{1}{\lambda}}\widehat{f})(\xi)$$

Proof.

$$\begin{aligned}
\widehat{\delta_\lambda f}(\xi) &= \int_{\mathbb{R}^n} \delta_\lambda f(x) e^{-2\pi i \xi x} dx \\
&= \int_{\mathbb{R}^n} f(\lambda x) e^{-2\pi i \xi x} dx \\
&= \frac{1}{\lambda^n} \int_{\mathbb{R}^n} f(y) e^{-2\pi i \frac{\xi}{\lambda} y} dy \\
&= \frac{1}{\lambda^n} \widehat{f}(\frac{\xi}{\lambda}) \\
&= \frac{1}{\lambda^n} (\delta_{\frac{1}{\lambda}} \widehat{f})(\xi)
\end{aligned}$$

■

Theorem 20.10 (Differentiation Dual to Multiplication)**.** *Suppose $f \in L^1(\mathbb{R}^n)$ with $-2\pi i x_k f(x) \in L^1(\mathbb{R}^n)$ with $1 \le k \le n$ and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Then $\hat{f}$ is differentiable and*

$$\frac{\partial}{\partial \xi_k} \hat{f}(\xi) = (-\widehat{2\pi i x_k} f)$$

Proof.

$$\begin{aligned}
(-\widehat{2\pi i x_k} f)(\xi) &= \int_{\mathbb{R}^n} (-2\pi i x_k f(x)) e^{-2\pi \sum_{i=1}^n \xi_i x_i} dx \\
&= \int_{\mathbb{R}^n} f(x) \cdot \frac{\partial}{\partial \xi_k} \left(e^{-2\pi i x \cdot x}\right) \; dx \\
&= \frac{\partial}{\partial \xi_k} \int_{\mathbb{R}^n} f(x) \left(e^{-2\pi i x \cdot x}\right) \; dx && \text{Dominated Convergence} \\
&= \frac{\partial}{\partial \xi_k} \hat{f}(\xi)
\end{aligned}$$

■

Remark 20.11. *Notice if $x_k f in L^1$, then $f \sim \frac{1}{|x_k|}$ as $x_k \to \infty$.*

Lemma 20.12. *Suppose f has compact support, then $\hat{f} \in C^\infty(\mathbb{R}^n)$.*

Lemma 20.13. *Suppose $f \in C^1(\mathbb{R}^n)$ and $\frac{\partial f}{\partial x_k} \in L^1(\mathbb{R}^n)$ and $\lim_{|x| \to \infty} f(x) = 0$. Then*

$$\widehat{\frac{\partial f}{\partial x_k}}(\xi) = 2\pi i \xi_k \hat{f}(\xi)$$

Proof.

$$\begin{aligned}
\widehat{\frac{\partial f}{\partial x_k}}(\xi) &= \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_k} e^{-2\pi i x \cdot \xi} \; dx \\
&= -\int_{\mathbb{R}^{n-1}} \left[\int_{\mathbb{R}} f(x) \frac{\partial}{\partial x_k} \left(e^{-2\pi i \xi \cdot x}\right) dx_k \right] d\hat{x_k} && \text{Integration by Parts} \\
&= 2\pi i \xi_k \int_{\mathbb{R}^n} f(x) e^{-2\pi i \xi \cdot x} \; dx \\
&= 2\pi i \xi_k \hat{f}(\xi)
\end{aligned}$$

■

Lemma 20.14. *$f, g \in L^1(\mathbb{R}^n)$, then*

$$\widehat{f * g}(\xi) = \widehat{f}(\xi)\widehat{g}(\xi)$$

Proof. By Young's Inequaulity, then $f * g \in L^1(\mathbb{R}^n)$. So we know we can take the transform of this. Observe:

$$\begin{aligned}
\widehat{f * g}(\xi) &= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(x-y)g(y)\ dy\right] e^{-2\pi i\xi\cdot x}\ dx \\
&= \int_{\mathbb{R}^n} f(x-y)dxe^{-2\pi i\xi\cdot(x-y)} \int_{\mathbb{R}^n} g(y)e^{-2\pi i\xi\cdot y)}\ dy \qquad \text{Fubini and Change of Variables} \\
&= \widehat{f}(\xi)\widehat{g}(\xi)
\end{aligned}$$

■

Lemma 20.15. *$f, g \in L^1$, then $\int_{\mathbb{R}^n} \hat{f}(x)g(x)\ dx = \int_{\mathbb{R}^n} f(\xi)\hat{g}(\xi)d\xi$*

Proof. $f, g \in L^1$, $\implies \hat{f}, \hat{g} \in L^\infty \implies f\hat{g}, \hat{f}g \in L^1$. Therefore, we can take the transform of the pairs.

$$\begin{aligned}
\int_{\mathbb{R}^n} \hat{f}(x)g(x)\ dx &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(\xi)e^{-2\pi i\xi\cdot x}d\xi g(x)\ dx \\
&= \int_{\mathbb{R}^n} f(\xi)\left(\int_{\mathbb{R}^n} g(x)\ e^{-2\pi i\xi\cdot x}\ dx\right)d\xi \qquad \text{Fubini} \\
&= \int_{\mathbb{R}^n} f(\xi)\hat{g}(\xi)d\xi
\end{aligned}$$

■

20.2 Examples of Fourier Transforms

Example 20.16. *Let $f(x) = \mathcal{X}_{[-a,a]} : \mathbb{R} \to \mathbb{R}$. Then*

$$\hat{f}(\xi) = \int_{-a}^{a} e^{-2\pi i\xi\cdot x}\ dx = \begin{cases} \frac{-e^{-2\pi i\xi a}+e^{2\pi i\xi a}}{2\pi i\xi} = \frac{\sin(2\pi\xi a)}{\pi\xi} & x \neq 0 \\ 2a & \xi = 0 \end{cases}$$

Notice, $\hat{f}(\xi) \in C^\infty(\mathbb{R}^n)$ but $\hat{f} \notin L^1(\mathbb{R}^n)$.

Question 20.17. *Is there such a function f in which $\hat{f} = f$?*

Example 20.18. *Yes. Consider $\phi(x) = e^{-\pi|x|^2}$.*

$$\begin{aligned}
\hat{\phi}(\xi) &= \int_{\mathbb{R}^n} e^{-\pi\sum_{k=1}^n x_k^2} e^{-2\pi i\sum_{k=1}^n \xi_k x_k}\ dx \\
&= \int_{\mathbb{R}^n} \prod_{k=1}^n e^{-\pi(x_k+i\xi_k)^2+\pi(i\xi_k)^2} dx_k \\
&= e^{-\pi|\xi|^2} \prod_{k=1}^n \underbrace{\int_{\mathbb{R}^n} e^{-\pi(x_k+i\xi_k)^2} dx_k}_{F(\xi_k)}
\end{aligned}$$

Now,

$$\begin{aligned}
\frac{dF}{d\xi_k}(\xi_k) &= \int_{\mathbb{R}^n} e^{-\pi(x_k+i\xi_k)^2}(-2\pi i(x_k+i\xi_k))dx_k \\
&= \int_{\mathbb{R}} i\frac{d}{dx_k}\left(e^{-\pi(x_k+i\xi_k)^2}\right) dx_k \\
&= 0
\end{aligned}$$

Therefore,

$$\begin{aligned}\hat{\phi}(\xi) &= e^{-\pi|\xi|^2}\prod_{k=1}^{n}\int_{\mathbb{R}^n} e^{-\pi(x_k+i\xi_k)^2}dx_k \\ &= e^{-\pi|\xi|^2}\prod_{k=1}^{n}\int_{\mathbb{R}^n} e^{-\pi x_k^2}dx_k \\ &= e^{-\pi|\xi|^2}\end{aligned}$$

20.3 Extending Fourier Transforms to L^2

Theorem 20.19. *We can extend the Fourier Transform to L^2 functions via the following statements:*

1. *Plancheral's Theorem: Suppose $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then $\hat{f} \in L^2(\mathbb{R}^n)$ and $\|f\|_{L^2} = \left\|\hat{f}\right\|_{L^2}$. Further, this shows the Fourier Transform is an isometry on L^2.*
2. *$f \to \hat{f}$ can be extended to L^2 functions.*
3. *Parseval's Identity to justify*
$$\langle f, g\rangle = \int f\overline{g}dx = \langle \hat{f}, \hat{g}\rangle$$

Proof. (1) implies (3) by the polarization identity.
(1) implies (2): If $f \in L^2(\mathbb{R}^n)$, we can find a sequence $\{f_j\} \subset L^1 \cap L^2(\mathbb{R}^n)$ such that

$$f_j \to f \text{ in } L^2$$

Therefore, $\{f_j\}$ is Cauchy, so

$$\limsup_{j,j\to\infty} \|f_j - f_k\|_{L^2} \to 0$$

But further, by Plancheral's Theorem:

$$\limsup_{j,j\to\infty} \left\|\hat{f}_j - \hat{f}_k\right\|_{L^2} \to 0$$

Therefore, $\hat{f} \to \hat{f}$ in L^2. Therefore, we can define

$$\hat{f} = \lim_{j\to\infty} \hat{f}_j$$

(1) To prove this, we need the fact that if we define

$$\phi(x) = e^{-\pi|x|^2}$$

and define $\lambda = \sqrt{4\pi\epsilon}$, then we see:

1. $\int_{\mathbb{R}^n} \hat{\phi}_\lambda(\xi)\ d\xi = 1$ where $\hat{\phi}_\lambda(\xi) = \frac{1}{\lambda^n}\hat{\phi}(\frac{\xi}{\lambda})$.
2. $\widehat{\widehat{\phi_\lambda}}(x) = \phi_\lambda(x)$

By these two facts, $\widehat{\phi_\lambda}$ can be used as an approximation to the identity. Now, since $f \in L^1$, we know $\hat{f} \in L^\infty$. Moreover,

$$\begin{aligned}
\int |\hat{f}|^2 d\xi &= \lim_{\epsilon \downarrow 0} \int \underbrace{\phi_\lambda(\xi)}_{\in L^1} |\hat{f}|^2 d\xi \\
&= \lim_{\epsilon \downarrow 0} \int \phi_\lambda(\xi) \hat{f} \cdot \overline{\hat{f}} d\xi \\
&= \lim_{\epsilon \downarrow 0} \int \left(\phi_\lambda(\xi) \int f(x) e^{-2\pi i \xi \cdot x} dx \overline{\int f(x) e^{-2\pi i \xi \cdot x} dx} \right) d\xi \\
&= \lim_{\epsilon \downarrow 0} \iiint \phi_\lambda(\xi) e^{-2\pi i \xi \cdot (x-y)} d\xi \ f(y) \ dy \ f(x) \ dx \\
&= \lim_{\epsilon \downarrow 0} \int \int \underbrace{\int \hat{\phi}_\lambda(x-y) \ f(y) \ dy}_{\hat{\phi}_\lambda * f} \ f(x) \ dx \\
&= \lim_{\epsilon \downarrow 0} \int \int \hat{\phi}_\lambda * f \ f(x) \ dx \\
&= \int \int \lim_{\epsilon \downarrow 0} \hat{\phi}_\lambda * f \ f(x) \ dx && \text{Dominated Convergence} \\
&= \lim_{\epsilon \downarrow 0} \int |f(x)|^2 \ dx
\end{aligned}$$

■

Remark 20.20. *Now we know the Fourier transform is a linear operator such that*

$$\begin{cases} \left\| \hat{f} \right\|_{L^\infty} \leq \|f\|_{L^1} \\ \left\| \hat{f} \right\|_{L^2} = \|f\|_{L^2} \end{cases}$$

Therefore, by the Riesz-Thorin Interpolation Theorem, we see, for $0 \leq \theta \leq 1$,:

$$\frac{1}{p} = \frac{1-\theta}{1} + \frac{\theta}{2} = 1 - \frac{\theta}{2}$$

$$\frac{1}{q} = \frac{1-\theta}{\infty} + \frac{\theta}{2} = \frac{\theta}{2}$$

Therefore, $1 \leq p \leq 2$. So we can extend the Transform for all $1 \leq p \leq 2$:

$$\left\| \hat{f} \right\|_{L^q} \leq C_{p,q,n} \|f\|_{L^p}$$

Remark 20.21. *To extend the Fourier Transform for $p > 2$, we need to resort to the theory of distributions.*

20.4 Fourier Inversion

Theorem 20.22 (Fourier Inversion)**.** *Suppose $f, \hat{f} \in L^1(\mathbb{R}^n)$. Then*

$$f(x) = \int \hat{f}(\xi) e^{2\pi i \xi \cdot x} d\xi \quad a.e.$$

Further, if f is continuous, then

$$f(x) = \int \hat{f}(\xi) e^{2\pi i \xi \cdot x} d\xi$$

Proof.

$$\begin{aligned}
f(x) &= \lim_{\epsilon \downarrow 0} \left(\widehat{\phi}_\lambda * f \right)(x) \\
&= \lim_{\epsilon \downarrow 0} \int f(x-y)\ \widehat{\phi_\lambda}(y)\ dy \\
&= \lim_{\epsilon \downarrow 0} \int f(x+y)\widehat{\phi_\lambda}(y)\ dy && \text{since } \widehat{\phi_\lambda}(-y) = \widehat{\phi_\lambda}(y) \\
&= \lim_{\epsilon \downarrow 0} \int \widehat{f(x+y)}\phi_\lambda(\xi)d\xi \\
&= \lim_{\epsilon \downarrow 0} \int e^{2\pi i \xi \cdot x}\widehat{f}(\xi)\phi_\lambda(\xi)d\xi \\
&\int \lim_{\epsilon \downarrow 0} e^{2\pi i \xi \cdot x}\widehat{f}(\xi)\phi_\lambda(\xi)d\xi && \text{By Dominated Convergence Theorem} \\
&\int \widehat{f}(\xi)e^{2\pi i \xi \cdot x}d\xi
\end{aligned}$$

■

Corollary 20.22.1 (Uniqueness)**.** *Suppose $f, g \in L^1(\mathbb{R}^n)$ and $\hat{f} = \hat{g}$. Then $f = g$ a.e.*

Proof. By the linearity of the Fourier Transform:

$$\widehat{f-g} = \hat{f} - \hat{g} = 0 \in L^1(\mathbb{R}^n)$$

Therefore, since $f, g \in L^1 \implies f - g \in L^1$, then we can apply the Inverse Fourier Transform

$$(f-g)(x) = \int_{\mathbb{R}^n} \widehat{f-g}e^{2\pi i \xi \cdot x}d\xi = 0$$

Therefore, $f = g$ a.e. ■

20.5 Application: The Heat Equation

Suppose we define $u(x,t)$ as temperature on a domain $\Omega \subset \mathbb{R}^n$. Further, we stipulate this function $u \in C^0(\overline{\mathbb{R}^n \times \mathbb{R}_{\geq 0}}) \cap C^2(\mathbb{R}^n \times \mathbb{R}_{\geq 0})$ satisfies the equations

$$\begin{cases} \partial_t u = \Delta u = \sum_{k=1}^n \frac{\partial^2 u}{\partial x_k^2} \\ u(x,0) = g(x) \end{cases}$$

This should model how the temperature distributes over time.
We can relax this problem by allowing $\Omega = \mathbb{R}^n$. Now the Fourier transform is able to transform a problem by changing the derivatives to multiplications.
Suppose $u \in L^1(\mathbb{R}^n)$ satisfies the above equations. Then

$$\hat{u}(x,t) = \int_{\mathbb{R}^n} u(x,t)e^{-2\pi i \xi \cdot x}dx$$

So

$$\partial_t \hat{u}(\xi, t) = \sum_{k=1}^n (2\pi i \xi_k)^2 \hat{u}$$

Therefore, we have reduced the problem to an ODE:

$$\begin{cases} \partial_t \hat{u}(\xi,t) = -2\pi i|\xi|^2 \hat{u} \\ \hat{u}(\xi, 0) = \hat{g}(\xi) \end{cases}$$

Therefore, integrating by t we get as a general solution of the defining ODE:

$$\hat{u}(\xi, t) = C(\xi)e^{-2\pi|\xi|^2 t}$$

Applying the initial condition, we see:

$$\hat{u}(\xi, 0) = C(\xi) = \hat{g}(\xi)$$

Therefore, we get a solution:

$$\hat{u}(\xi, t) = \hat{g}(\xi)e^{-2\pi|\xi|^2 t}$$

Now we reach the issue of returning this from the frequency domain back into the space domain. If we define $\widehat{H_t} := e^{-2\pi|\xi|^2 t}$, then we see:

$$\hat{u}(\xi, t) = \hat{g}(\xi)\widehat{H_t}(\xi)$$

And therefore,

$$u(x, t) = (g * H_t)(x, t)$$

Now, taking the inverse of $\widehat{H_t}$ we see:

$$H_t = (4\pi\epsilon)^{\frac{n}{2}}\phi_{\sqrt{4\pi\epsilon}}(x) = \left(\frac{1}{2t}\right)^{n/2}\phi_{\sqrt{\frac{n}{2}}} = \frac{1}{(2t)^{n/2}}e^{-\pi\frac{|x|^2}{2t}}$$

where $\epsilon = \frac{1}{8\pi t}$. This function is known as the Heat Kernel. Therefore,

$$u(x, t) = \frac{1}{(2t)^{n/2}}\int_{\mathbb{R}^n} g(x - y)\frac{1}{(2t)^{n/2}}e^{-\pi\frac{|y|^2}{2t}}\,dy$$

Remark 20.23. *We see the norm of our solution is controlled:*

$$\sup_{x\in\mathbb{R}^n}|u(x, t)| \le \frac{1}{\sqrt{2t}^n}\int |g|\ dy$$

And we can see:

$$\lim_{t\to\infty}|u(x, t)| = 0$$

Physically, this tells us the temperature dissipates to zero.

Remark 20.24. *Further, we can rewrite this formula as*

$$u(x, t) = \frac{1}{\sqrt{2t}^n}\int_{\mathbb{R}^n} g(y)e^{\frac{-\pi(x-y)^2}{2t}}\ dt$$

This interpretation provides more physical insights. If g is nonzero somewhere, this equation tells us that u will be instantaneously nonzero. Moreover, this solution will be continuous, no matter how rough the initial condition was.

20.6 The Schwartz Class

Definition 20.25. *The Space of Schwartz Functions, denoted $S(\mathbb{R}^n)$, is defined as:*

$$\mathcal{S}(\mathbb{R}^n) := \{u \in C^\infty(\mathbb{R}^n) : \|u\|_{\alpha,\beta} = \sup_{x\in\mathbb{R}^n}|x^\alpha D^\beta u| < \infty\ \forall\alpha, \beta \in \mathbb{Z}^n_{\ge 0}\}$$

where we equip this space with a semi-norm:

$$[f]_{\alpha,\beta} := \sup_{x\in\mathbb{R}^n}|x^\alpha D^\beta u|$$

In essence, we should interpret this as both the function and its derivatives must decay faster than all polynomials at infinity. It's sometimes referred to as the space of rapidly decreasing functions. As a result, we draw several immediate conclusions:

Remark 20.26. $C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$

Remark 20.27. $e^{c|x|^2}\mathcal{S}(\mathbb{R}^n)$ *provided* $c > 0$.

Remark 20.28. *$\mathcal{S}(\mathbb{R}^n)$ is a vector space and an algebra. To see this, Suppose $f, g \in \mathcal{S}(\mathbb{R}^n)$ and*

$$\left\|x^\alpha D^\beta(fg)\right\|_\infty = \left\|x^\alpha \sum c_{\beta,k} D^k f^{\beta-k} g\right\|_{L^\infty} \leq \sum c_{\beta,k} \left\|x^\alpha D^k f\right\|_{L^\infty} \left\|x^\alpha D^k f\right\|_{L^\infty} < \infty$$

Definition 20.29. *Given $\{f_n\} \subset \mathcal{S}(\mathbb{R}^n), f \in \mathcal{S}(\mathbb{R}^n)$, we say $f_n \xrightarrow{\mathcal{S}} f$ provided for all α, β,*

$$[f_n - f]_{\alpha,\beta} \to 0$$

This seminorm inspires us to endow the space with a metric.

Definition 20.30. *$\mathcal{S}(\mathbb{R}^n)$ is metrizable with metric:*

$$d(f,g) := \sum_{|\alpha|,|\beta| \geq 0} 2^{-|a|-|b|} \frac{[f-g]_{\alpha,\beta}}{1 + [f-g]_{\alpha,\beta}}$$

Lemma 20.31. *$(\mathcal{S}(\mathbb{R}^n), d)$ is a complete metric space.*

Lemma 20.32. *Given $f \in \mathcal{S}(\mathbb{R}^n)$, then $p(x)f \in \mathcal{S}(\mathbb{R}^n)$ for all polynomials $p(x)$.*

Lemma 20.33. *$C_c^\infty(\mathbb{R}^n)$ is dense in $\mathcal{S}(\mathbb{R}^n)$.*

Lemma 20.34. *If $f \in \mathcal{S}(\mathbb{R}^n)$, then $f \in L^p(\mathbb{R}^n)$ for all $1 \leq p \leq \infty$*

Proof.

$$\begin{aligned}
\int |f|^p \, dx &= \int \frac{|f(1+|x|)^k|^p}{(1+|x|)^{kp}} \\
&\leq \left\|f(1+|x|)^k\right\|_{L^\infty}^p \int \frac{1}{(1+|x|)^{kp}} \\
&\leq \left\|f(1+|x|)^k\right\|_{L^\infty}^p \int \frac{r^{n-1}}{(1+r)^{kp}} \\
&< \infty & kp > n
\end{aligned}$$

■

Now we can extend the Fourier transform onto this space:

Theorem 20.35. *$F : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is a continuous map which follow the properties:*

- $\mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$
- $F(\mathcal{S}(\mathbb{R}^n)) \subset L^\infty(\mathbb{R}^n)$

Proof. We need to demonstrate

$$\left\|\xi^\alpha D^\beta \hat{f}(\xi)\right\|_{L^\infty} = \left\|(-2pii)^{|\alpha|+|\beta|} \widehat{x^\beta D^\alpha f}\right\|_{L^\infty} < \infty$$

■

Lemma 20.36. *If $f \in \mathcal{S}(\mathbb{R}^n)$, then*

$$f = \int \hat{f}(\xi) e^{2\pi i \xi \cdot x} \, dx$$

Proof. If $f \in \mathcal{S}(\mathbb{R}^n)$, then $\hat{f} \in \mathcal{S}(\mathbb{R}^n) \implies f, \hat{f} \in L^1 \implies f = F^{-1}(\hat{f})$ a.e. But since $f \in C^\infty$, then $f = F^{-1}(\hat{f})$ ■

Lemma 20.37. *The Fourier Transform on $F : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is a bijection!*

20.7 Tempered Distributions

Definition 20.38. *A Tempered Distribution u is a bounded, linear functional on $\mathcal{S}(\mathbb{R}^n)$:*

$$u : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$$

provided u is linear and $\langle u, \phi_n \rangle \to 0$ for all ϕ_n such that $\phi_n \xrightarrow{\mathcal{S}} 0$.

Remark 20.39. *A distribution is a bounded, linear functional on $C_c^\infty(\mathbb{R}^n)$.*

Definition 20.40. *The Space of Tempered Distributions is denoted $\mathcal{S}'(\mathbb{R}^n)$. Further, $u \in \mathcal{S}'(\mathbb{R}^n)$, provided u is a linear functions on $\mathcal{S}(\mathbb{R}^n)$ and there exists constant C, N, M such that*

$$|\langle u, \phi \rangle| \leq C \sum_{|\alpha| \leq N, |\beta| \leq M} [\phi]_{\alpha,\beta}$$

for all $\phi \in \mathcal{S}(\mathbb{R}^n)$.

Theorem 20.41 (Structural Theorem of $\mathcal{S}'(\mathbb{R}^n)$)**.** *Suppose $u \in \mathcal{S}'(\mathbb{R}^n)$. Then there exist finitely many bounded continuous functions*

$$u_{\alpha,\beta} : \mathbb{R}^n \to \mathbb{C}$$

with $|\alpha| \leq N, |\beta| \leq M$ such that

$$u = \sum_{|\alpha| \leq N, |\beta| \leq M} x^\alpha D^\beta u_{\alpha,\beta}$$

Lemma 20.42. *$L^p(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$ for $1 \leq p \leq \infty$.*

Proof. To show this, we need to demonstrate how each L^p function pairs with a function in the Schwartz space. Observe, given $u \in L^p(\mathbb{R}^n)$,

$$|\langle u, \phi \rangle| = |\int u \cdot \phi| \leq \|u\|_{L^p} \|\phi\|_{L^q} < \infty$$

for all $\phi \in \mathcal{S}(\mathbb{R}^n)$. Moreover, the continuity of u is obvious. ∎

Example 20.43. *The dirac delta function $\delta \in \mathcal{S}'(\mathbb{R}^n)$ since*

$$\langle \delta, \phi \rangle \phi(0)$$

and since $\phi_n \xrightarrow{\mathcal{S}} 0 \implies \phi_n(0) \to 0$, then

$$\langle \delta, \phi \rangle \to 0$$

We refer to this distribution as the dirac delta distribution.

Lemma 20.44. *If $f \in \mathcal{S}'(\mathbb{R}^n)$ then $D^\alpha f \in \mathcal{S}'(\mathbb{R}^n)$ for all α.*

Proof. Observe,

$$\langle D^\alpha f, \phi \rangle := (-1)^\alpha \langle f, D^\alpha \phi \rangle$$

Since $f \in \mathcal{S}'(\mathbb{R}^n)$, then $(-1)^\alpha \langle f, D^\alpha \phi \rangle$ is a bounded linear functional. ∎

Example 20.45.

$$\begin{aligned} \langle D^\alpha \delta, \phi \rangle &= (-1)^\alpha \langle \delta, D^\alpha \phi \rangle \\ &= (-1)^\alpha \langle \delta, D^\alpha \phi \rangle \\ &= (-1)^\alpha D^\alpha \phi(0) \end{aligned}$$

Example 20.46. *Define the Heaviside function:*

$$H(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Then we see:

$$\begin{aligned} \langle H', \phi \rangle &= (-1)\langle H, \phi' \rangle \\ &= (-1)\int_0^\infty \phi'(x)\ dx \\ &= \phi(0) \end{aligned}$$

Therefore, $H' = \delta$!

Example 20.47. *Consider the function*

$$f(x) = \frac{1}{|x|^{n+\alpha}}$$

for any $\alpha \in [0,1)$. Then for any $\phi \in \mathcal{S}(\mathbb{R}^n)$

$$\langle\ p.v.\ \frac{1}{|x|^{n+\alpha}}, \phi\rangle = \lim_{\epsilon \downarrow 0} \int_{|x|>\epsilon} \frac{\phi(x)-\phi(0)}{|x|^{n+\alpha}}\ dx = \lim_{\epsilon \downarrow 0} \underbrace{\int_{1>|x|>\epsilon} \frac{\phi(x)-\phi(0)}{|x|^{n+\alpha}}\ dx}_{I} + \underbrace{\int_{|x|>1} \frac{\phi(x)-\phi(0)}{|x|^{n+\alpha}}\ dx}_{II}$$

We see we can bound the integrals:

$$II \leq \int_{|x|>1} \frac{2\left\|\phi\right\|_{L^\infty}}{|x|^{n+\alpha}} dx < \infty$$

and we apply the mean value theorem to control the first integral:

$$I = \int_{1>|x|>\epsilon} \frac{\phi(x)-\phi(0)}{|x|^{n+\alpha}}\ dx \leq |x| \left\|\nabla\phi\right\|_{L^\infty} \leq \left\|\nabla\phi\right\|_{L^\infty}$$

Further, this inner product converges to zero as $\epsilon \downarrow 0$.

Definition 20.48. *For any $f \in \mathcal{S}'(\mathbb{R}^n)$, then*

$$\langle \hat{f}, \phi \rangle := \langle f, \hat{\phi} \rangle$$

for any $\phi \in \mathcal{S}(\mathbb{R}^n)$

Remark 20.49. *Notice, $\langle f, \hat{\phi} \rangle$ is a bounded linear function. So in fact, $\hat{f}$ is a tempered distribution.*

Example 20.50.

$$\langle \hat{\delta}, \phi \rangle = \langle \delta, \hat{\phi} \rangle = \hat{\phi}(0) = \int \phi(x)\ dx = \langle 1, \phi \rangle$$

Therefore, $\hat{\delta} = 1$.

Definition 20.51. *The translation operator $\tau_h : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is*

$$T_h\phi(x) = \phi(x-h)$$

Definition 20.52. *The relection operator $R : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is*

$$Rf(x) = f(-x)$$

Lemma 20.53. *For any tempered distribution $T \in \mathcal{S}'(\mathbb{R}^n)$,*

$$\langle \tau_h T, \phi \rangle := \langle T, \tau_{-h}\phi \rangle$$

for all $\phi \in \mathcal{S}(\mathbb{R}^n)$.

Proof.

$$\int (\tau_h f)\phi = \int f(x-h)\phi(x) = \int f \cdot (\tau_{-h}\phi)(x)\ dx$$

■

Lemma 20.54. *For any tempered distribution $T \in \mathcal{S}'(\mathbb{R}^n)$,*

$$\langle RT, \phi \rangle := \langle T, R\phi \rangle$$

for all $\phi \in \mathcal{S}(\mathbb{R}^n)$.

Proof.

$$\int (Rf)\phi = \int f(-x)\phi(x) = \int f \cdot (R\phi)(x)\ dx$$

■

Lemma 20.55 (Convolution)**.** *For any $\phi \in \mathcal{S}(\mathbb{R}^n)$ and $u \in \mathcal{S}'(\mathbb{R}^n)$, then*

$$(\phi * u)(x) = \langle u, R\tau_x\phi \rangle$$

Proof.

$$\begin{aligned}
(\phi * u)(x) &= \int u(x-y)\phi(y)\ dy \\
&= \int \phi(x-y)u(y)\ dy \\
&= \int R\phi(y-x)u(y)\ dy \\
&= \int (R\tau_x\phi)(y)u(y)\ dy \\
&= \langle u, R\tau_x\phi \rangle
\end{aligned}$$

■

Example 20.56.

$$(\phi * \delta)(x) = \langle \delta, R\tau_x\phi = R\tau_x\phi(x) = R\phi(-x) = \phi(x)$$

Now, we've translated all of the properties of the Fourier Transform into the extended realm of tempered distributions. Therefore, we can now understand the Fourier Transform in this more generalized setting.

20.8 Application: Fractional Derivative $s > 0$

Given $f : \mathbb{R} \to \mathbb{R}$, then we know

$$\hat{f'} = (2\pi i\xi)\hat{f}(\xi)$$

This translates differentiation into multiplication.

Question 20.57. *Is it possible to multiply by a fraction and therefore have a fractional derivative? That is,*

$$\widehat{\frac{d^{(s)}}{dx^{(s)}}} f = (2\pi i\xi)^s\hat{f}(\xi)?$$

Definition 20.58. *Given $s \in \mathbb{R}$, define the space*

$$H^s(\mathbb{R}^n) := \{f \in \mathcal{S}(\mathbb{R}^n) : (\sqrt{1+|\xi|^2})^s \hat{f}(\xi) \in L^2(\mathbb{R})\}$$

- $s > 0 \implies$ *fractional differentiation.*
- $s < 0 \implies$ *fractional integration.*

We equip this space with the norm:

$$\|f\|_{H^s(\mathbb{R}^n)} := \left\|(\sqrt{1+|\xi|^2})^s \hat{f}(\xi)\right\|_{L^2(\mathbb{R}^n)}$$

Lemma 20.59. *If $s \in \mathbb{Z}_{\geq 0}$, then*

$$\|f\|_{H^s(\mathbb{R}^n)} \text{ is equivalent to } \|f\|_{W^{s,2}(\mathbb{R}^n)}$$

That is, they generate the same topology.

Proof. Case: $s = 1$, then

$$\begin{aligned}
\|f\|^2_{H^1(\mathbb{R}^n)} &= \int [\sqrt{1+|\xi|^2}|\hat{f}|]^2 \\
&= \int |\hat{f}|^2 d\xi + \int |\xi \hat{f}|^2 d\xi \\
&= \int |f|^2\, dx + \frac{1}{2\pi}\int |\widehat{Df}|^2\, d\xi && \text{Plancherel} \\
&= \int |f|^2\, dx + \frac{1}{2\pi}\int |Df|^2\, dx
\end{aligned}$$

■

Theorem 20.60 (Trace - Improved!)**.** *Suppose $s > \frac{1}{2}$. Let $T : C_c^\infty(\mathbb{R}^n) \to C_c^\infty(\mathbb{R}^{n-1})$ is defined as*

$$(Tf)(x_1, x_2, \ldots, x_{n-1}, x_n) = f(x_1, \ldots, x_{n-1}, 0)$$

Then T extends to a bounded linear operator from $H^s(\mathbb{R}^n) \to H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})$.

Before we prove this, recall the fact that for a given function $f : \mathbb{R} \to \mathbb{R}$, then

$$f(0) = \int_{\mathbb{R}} \hat{f}(\xi)\, d\xi$$

Proof. We will define:

$$\tilde{f}(\xi') := \int_{\mathbb{R}^{n-1}} f(\underbrace{x_1, \ldots, x_{n-1}}_{x'}, 0) e^{-2\pi i \xi' \cdot x'} dx'$$

with $\xi' = (\xi_1, \ldots, \xi_{n-1}) \in \mathbb{R}^{n-1}$. Now, we need to show that $\|Tf\|_{H^{s-1/2}(\mathbb{R}^{n-1})} \leq C_s \|f\|_{H^s(\mathbb{R}^n)}$ for some constant. That is,

$$\|Tf\|_{H^{s-1/2}(\mathbb{R}^{n-1})} = \int_{\mathbb{R}^{n-1}} (1+|\xi'|^2)^{s-1/2} |\tilde{f}(\xi')|^2 d\xi' \leq C_s \int_{\mathbb{R}^n} (1+|\xi|^2)^s |\hat{f}|^2 d\xi$$

We will prove this by proving a claim first:
Claim: $\int_{\mathbb{R}} \hat{f}(\xi'.\xi_n) d\xi_n = \tilde{f}(\xi')$

$$\begin{aligned}
\int_{\mathbb{R}} \hat{f}(\xi'.\xi_n)d\xi_n &= \int_{\mathbb{R}}\int_{\mathbb{R}^n} f(x)e^{-2\pi i(\xi'\cdot x'+\xi_n x_n)}dx'\ dx_n d\xi_n \\
&= \int_{\mathbb{R}^{n-1}}\int_{\mathbb{R}}\int_{\mathbb{R}} f(x)e^{-2\pi i\xi_n x_n}dx_n d\xi_n\ e^{-2\pi i\xi'\cdot x'}dx' \\
&= \int_{\mathbb{R}^{n-1}}\int_{\mathbb{R}} \hat{f}^{x_m}(x',\xi_n)d\xi_n dx' \\
&= \int_{\mathbb{R}^{n-1}} f(x',0)e^{-2\pi i\xi'\cdot x'}dx' \\
&= \tilde{f}(\xi')
\end{aligned}$$

We can now use this claim to show:

$$\begin{aligned}
|\tilde{f}(\xi')|^2 &= \left[\int_{\mathbb{R}} \hat{f}(\xi',\xi_n)d\xi_n\right]^2 \\
&= \left[\int_{\mathbb{R}} \frac{(1+|\xi|^2)^{s/2}}{(1+|\xi|^2)^{s/2}}\hat{f}(\xi',\xi_n)d\xi_n\right]^2 \\
&\leq \int_{\mathbb{R}}(1+|\xi|^2)^s|\hat{f}(\xi',\xi_n)|^2 d\xi_n \int_{\mathbb{R}} \frac{d\xi_n}{(1+|\xi|^2)^s}
\end{aligned}$$

Observe,

$$\int_{\mathbb{R}} \frac{d\xi_n}{(1+|\xi|^2)^s} = \int \frac{\sqrt{1+|\xi'|^2}dt}{[(1+|\xi'|^2)(1t^2)]^s} = (1+|\xi'|^2)^{1/2-s}\int \frac{dt}{(1+t^2)^s} = C_s(1+|\xi'|^2)^{1/2-s}$$

Therefore, we see:

$$\begin{aligned}
&|\tilde{f}(\xi')|^2 \leq C_s(1+|\xi'|^2)^{1/2-s}\int_{\mathbb{R}}(1+|\xi|^2)^s|\hat{f}(\xi',\xi_n)|^2 d\xi_n \\
\iff &|\tilde{f}(\xi')|^2(1+|\xi'|^2)^{s-1/2} \leq C_s\int_{\mathbb{R}}(1+|\xi|^2)^s|\hat{f}(\xi',\xi_n)|^2 d\xi_n \\
\iff &\int_{\mathbb{R}^{n-1}}|\tilde{f}(\xi')|^2(1+|\xi'|^2)^{s-1/2}d\xi' \leq C_s\int_{\mathbb{R}^{n-1}}\int_{\mathbb{R}}(1+|\xi|^2)^s|\hat{f}(\xi',\xi_n)|^2 d\xi_n\ d\xi' = C_s\|f\|_{H^s(\mathbb{R}^n)}
\end{aligned}$$

∎

Theorem 20.61 (Poisson Summation Formula)**.** *Suppose* $f \in \mathcal{S}(\mathbb{R})$. *Then*

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{k\in\mathbb{Z}} \hat{f}(k)$$

The proof is the result if a nice trick called *periodizing the function.*

Proof. Define

$$F(x) := \sum_{n\in\mathbb{Z}} f(x+n) \text{ with } F(x+1) = F(x)$$

Further, this function admits a Fourier series:

$$F(x) = \sum_{n\in\mathbb{Z}} a_n e^{-2\pi i n x}$$

where $a_n := \int_0^1 F(x)e^{-2\pi inx}\, dx$. Therefore,

$$\begin{aligned} a_n &= \int_0^1 F(x)e^{-2\pi inx}\, dx \\ &= \int_0^1 \sum_{m\in\mathbb{Z}} f(x+n)e^{-2\pi imx}\, dx \\ &= \sum_{m\in\mathbb{Z}} \int_0^1 f(x+n)e^{-2\pi imx}\, dx \\ &= \sum_{m\in\mathbb{Z}} \int_0^1 f(x+n)e^{-2\pi im(x+n)}\, d(x+n) \\ &= \int_{-\infty}^{\infty} f(t)e^{-2\pi imt}\, dt \\ &= \hat{f}(m) \end{aligned}$$

Therefore,

$$F(x) = \sum_{m\in\mathbb{Z}} \hat{f}(m)e^{-2\pi imx}$$

Therefore,

$$\sum_{n\in\mathbb{Z}} f(x+n) = \sum_{m\in\mathbb{Z}} \hat{f}(m)e^{-2\pi imx}$$

Now, setting $x = 0$, we arrive at:

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{m\in\mathbb{Z}} \hat{f}(m)$$

■

20.9 Application: Heisenberg Uncertainty Principle

Theorem 20.62 (Heisenberg Uncertainty Principle). *If $f, xf, \xi\hat{f}(\xi) \in L^2(\mathbb{R})$, then*

$$\int_{\mathbb{R}} |xf|^2\, dx \int_{\mathbb{R}} |\xi\hat{f}(\xi)|^2 d\xi \geq \frac{1}{16\pi^2}\left(\int_{\mathbb{R}} |f|^2\, dx\right)^2$$

To interpret this, we can think of the left hand side integrals as measures of support. This principle tells us that you cannot simultaneously make both of these small. Similarly, the principle translates to physical meaning to imply both velocity and position cannot be known simultaneously.

Proof. Assume $f \in \mathcal{S}(\mathbb{R})$. Then we see:

$$\begin{aligned}
\int_{\mathbb{R}} |xf|^2 \, dx \int_{\mathbb{R}} |\xi \hat{f}(\xi)|^2 d\xi &= \int_{\mathbb{R}} |xf|^2 \, dx \int_{\mathbb{R}} |\xi \frac{1}{2\pi} \widehat{f'(\xi)}|^2 d\xi \\
&= \frac{1}{4\pi^2} \int_{\mathbb{R}} |xf|^2 \, dx \int_{\mathbb{R}} |\widehat{f'(\xi)}|^2 d\xi \\
&= \frac{1}{4\pi^2} \int_{\mathbb{R}} |xf|^2 \, dx \int_{\mathbb{R}} |f'(x)|^2 dx && \text{Plancheral} \\
&\geq \frac{1}{4\pi^2} \int_{\mathbb{R}} |xf(x)f'(x)|^2 \, dx && \text{Cauhcy Schwarz} \\
&\geq \frac{1}{4\pi^2} \left(\int_{\mathbb{R}} |xf(x)f'(x) \, dx \right)^2 \\
&= \frac{1}{16\pi^2} \left(\int_{\mathbb{R}} x \frac{d}{dx} (f(x))^2 \right)^2 \\
&= \frac{1}{16\pi^2} \left(\int_{\mathbb{R}} f^2(x) \, dx \right)^2 && \text{Integration by Parts}
\end{aligned}$$

■

This motivates representing a function by its derivative.

Example 20.63. *Consider the space $\mathbb{R}$. Suppose we are given $u \in C_c^1(\mathbb{R})$. Then*

$$u(x) = \int_{-\infty}^{x} u'(y) \, dy$$

Similarly,

$$u(x) = -\int_{x}^{\infty} u'(y) \, dy$$

Adding these two representations:

$$2u(x) = \int_{-\infty}^{x} u'(y) \, dy - \int_{x}^{\infty} u'(y) \, dy = \int_{-\infty}^{\infty} u'(y) \frac{x-y}{|x-y|} dy$$

Therefore,

$$u(x) = \frac{1}{2} \int_{-\infty}^{\infty} u'(y) \frac{x-y}{|x-y|} dy$$

Example 20.64. *Consider the space $\mathbb{R}^n$. Fix a unit vector $\vec{e} \in \partial B_1(0)$. Then we can write:*

$$u(x) = -\int_{0}^{\infty} \frac{\partial}{\partial r} u(x + r\vec{e}) dr = -\int_{0}^{\infty} \nabla u(x + r\vec{e}) \cdot e dr$$

On the other hand, we'll let $\omega_{n-1} := |\partial B_1(0)|$. Then we see:

$$
\begin{aligned}
\omega_{n-1} u(x) &= \int_{\partial B_1(0)} dS_e \\
&= -\int_0^\infty \nabla u(x + r\vec{e}) \cdot \vec{e} dr dS_e \\
&= -\int_0^\infty \int_{\partial B_1(0)} \nabla u(x + r\vec{e}) \cdot \vec{e} dS_e dr && \textit{Fubini} \\
&= -\int_0^\infty \int_{\partial B_1(0)} \nabla u(x + y) \frac{y}{r} \frac{dS_y}{r^{n-1}} dr && y = r\vec{e} \\
&= \int_0^\infty \int_{\partial B_r(0)} \nabla u(x + y) \frac{y}{|y|^n} dS_y dr \\
&= \int_{\mathbb{R}^n} \nabla u(z) \frac{x - z}{|x - z|^n} dz
\end{aligned}
$$

Therefore,

$$
u(x) = \frac{1}{\omega_{n-1}} \int_{\mathbb{R}^n} \nabla u(z) \frac{x - z}{|x - z|^n} dz
$$

This leads directly to the following lemma:

Lemma 20.65 (Riesz Potential)**.**

$$
|u(x)| \leq \frac{1}{\omega_{n-1}} \int_{\mathbb{R}^{n-1}} \frac{|\nabla u(y)|}{|x - y|^{n-1}} dy
$$

Definition 20.66. *For $0 < \alpha < n$, we define the generalized Riesz Potential*

$$
(I_\alpha f)(x) := \int_{\mathbb{R}^n} \frac{f(x)}{|x - y|^{n-\alpha}} dy
$$

Remark 20.67. *We can rewrite the previous lemma by*

$$
|u(x)| \leq \frac{1}{\omega_{n-1}} I_1(|\nabla u|)(x)
$$

Theorem 20.68 (Sobolev Inequality for Riesz Potential)**.** *Suppose $\alpha > 1, p > 1$ and $\alpha p < n$. Then there exists some constant $C_{n,p,\alpha}$, such that for any $f \in L^p(\mathbb{R}^n)$, then*

$$
\|I_\alpha f\|_{L^{p*}(\mathbb{R}^n)} \leq C_{n,p,\alpha} \|f\|_{L^p}
$$

where $p^* := \frac{np}{n-\alpha p}$.

Before we can prove this theorem, we need a couple helpful lemmas.

Lemma 20.69. *For $0 < \alpha < n$, there exists a constant $C_{\alpha,n}$ such that*

$$
\int_{B_r(x)} \frac{f(y)}{|x - y|^{n-\alpha}} dy \leq C r^\alpha \mu f(x)
$$

Proof. Fix $x \in \mathbb{R}^n$. We can decompose a ball surround x into Annuli

$$
B_r(x) := \bigcup_{j=0} B_{\frac{r}{2^j}}(x) \setminus B_{\frac{r}{2^{j+1}}}(x)
$$

Therefore,

$$\begin{aligned}
\int_{B_r(x)} \frac{f(y)}{|x-y|^{n-\alpha}} dy &= \sum_{j=0}^{\infty} \int_{B_{\frac{r}{2^j}}(x) \setminus B_{\frac{r}{2^{j+1}}}(x)} \frac{f(y)}{|x-y|^{n-\alpha}} dy \\
&\leq \sum_{j=0}^{\infty} \frac{1}{(\frac{r}{2^{j+1}})^{n-\alpha}} \int_{B_{\frac{r}{2^j}}(x)} |f(y)| dy \\
&= \sum_{j=0}^{\infty} \left(\frac{r}{2^{j+1}}\right)^{\alpha} \left(\frac{1}{2}\right)^{n} \fint_{B_{\frac{r}{2^j}}(x)} |f(y)| dy \\
&\leq C r^{\alpha} \mu f(x)
\end{aligned}$$

■

Lemma 20.70. *For $0 < \alpha p < n$, there exists a constant $C_{\alpha,n}$ such that*

$$\int_{\mathbb{R}^n \setminus B_r(x)} \frac{f(y)}{|x-y|^{n-\alpha}} dy \leq C r^{\alpha - \frac{n}{p}} \|f\|_{L^p(\mathbb{R}^n)}$$

Proof. As a consequence of Hölder's inequality:

$$\int_{\mathbb{R}^n \setminus B_r(x)} \frac{f(y)}{|x-y|^{n-\alpha}} dy \leq \|f\|_{L^p(\mathbb{R}^n)} \left\| \frac{1}{|x-y|^{n-\alpha}} \right\|_{L^q(\mathbb{R}^n \setminus B_r(x))}$$

Evaluating the second norm, we see upon converting to polar coordinates:

$$\begin{aligned}
\left\| \frac{1}{|x-y|^{n-\alpha}} \right\|_{L^q(\mathbb{R}^n \setminus B_r(x))} &= \omega_{n-1} \int_r^{\infty} \frac{1}{\rho^{n-\alpha}} \rho^{n-1} d\rho \\
&= \omega_{n-1} \int_r^{\infty} \rho^{(\alpha-n)\frac{p}{p-1}+n-1} d\rho \\
&= \omega_{n-1} \left[\frac{\rho^{\frac{(\alpha-n)p}{p-1}+n}}{\frac{(\alpha-n)p}{p-1}+n} \right]_r^{\infty} \\
&= C \cdot r^{\frac{\alpha p - n}{p-1}} \\
&= C r^{\alpha - \frac{n}{p}}
\end{aligned}$$

■

Proof. Of Sobolev Inequality
Observe:

$$\begin{aligned}
I_\alpha f(x) &= \int_{\mathbb{R}^n} \frac{f(x)}{|x-y|^{n-\alpha}} dy \\
&= \int_{\mathbb{R}^n \setminus B_r(x)} \frac{f(x)}{|x-y|^{n-\alpha}} dy + \int_{B_r(x)} \frac{f(x)}{|x-y|^{n-\alpha}} dy
\end{aligned}$$

For a convenient choice of r. Specifically,

$$r := \left(\frac{\mu f(x)}{\|f\|_{L^p}} \right)^{-p/n}$$

Then we can apply the previous lemmas to see:

$$I_\alpha f(x) \leq C r^{\alpha - n/p} \|f\|_{L^p} + C' r^{\alpha} \mu f(x) = C \mu f^{1 - \frac{\alpha p}{n}} \|f\|_{L^p}^{\frac{\alpha p}{n}}$$

Therefore, raising both sides and taking the $p*$ norm, we get:

$$\|I_\alpha f\|_{L^{p*}}^{p*} \leq C \, \|f\|^{\frac{\alpha p}{n} p*} \int \mu f^{p*(1-\frac{\alpha p}{n})} dx \leq C \, \|f\|_{L^p}^{\frac{\alpha p}{n} p*} \, \|f\|_{L^p}^{p}$$

Therefore,

$$\|I_\alpha f\|_{L^{p*}} \leq C \, \|f\|_{L^p}^{\alpha \frac{p}{n} + \frac{p}{p*}} = \|f\|_{L^p}$$

■

Corollary 20.70.1. *This theorem implies that for $f \in W_0^{1,p}(\mathbb{R}^n)$*

$$\|f\|_{L^{p*}(\mathbb{R}^n)} \leq \|Df\|_{L^p(\mathbb{R}^n)}$$

Proof. Of Corollary
We can write,

$$f(x) = \frac{1}{\omega_{n-1}} \int_{\mathbb{R}^n} \frac{\nabla f(y) \cdot (x-y)}{|x-y|^n} dy$$

Therefore, applying the Sobolev Inequality with $\alpha = 1$

$$\|f\|_{L^{p*}(\mathbb{R}^n)} \leq \|I_1(|\nabla f|)\|_{L^{p*}} \leq C \, \|\nabla f\|_{L^p(\mathbb{R}^n)}$$

■

21 Additional Topics

21.1 John-Nirenberg Inequality

Theorem 21.1 (John-Nirenberg Inequality)**.** *Suppose $\|f\|_{BMO} < \infty$, the for all $Q \subset \mathbb{R}^n$,*

$$|\{x \in Q : |f(x) - f_Q| > \lambda\}| \leq C_1 |Q| \exp\left(\frac{-C_2\lambda}{\|f\|_{BMO}}\right)$$

with $C_1 > 0, C_2 > 0$ depending only on n.

Remark 21.2. $L^\infty(\mathbb{R}^n) \not\subset BMO(\mathbb{R}^n)$.

Question 21.3. *Suppose we have an operator T such that when restricted to $T : L^2 \to L^2$ and $T : L^\infty \to BMO$ are bounded, does this imply for any $2 < p < \infty$,*

$$T : L^p \to L^p$$

a bounded operator?

As it turns out, the BMO space serves as a nice substitute for L^∞.

Proof. (Of Theorem)
Fix $Q \subset \mathbb{R}^n$. Define

$$F(\lambda) := \sup_{Q' \subset Q} \frac{1}{|Q'|}\{x \in Q'|f(x) - f_Q| > \lambda\}$$

Equivalently, $F(\lambda)$ is the smallest number such that

$$|\{x \in Q' : |f(x) - f_Q|\}| \leq |Q'|F(\lambda)$$

We will now use iteration in order to bound $F(\lambda)$. Recall the Calderón-Zygmund Decomposition Theorem, which tells us:

- $|f - f_Q| \in L^1(Q)$
- There exists $\{Q_i\}$ pairwise disjoint cubes such that $Q = \bigcup Q_i$
  1. $|f - f_Q| \leq 2^n$ for almost every $x \in Q^c$
  2. $|\bigcup Q_i| \leq \frac{\|f - f_Q\|_{L^1(Q)}}{2^n} \leq \frac{|Q|\|f\|_{BMO}}{2^n}$
  3. $2^n < \frac{1}{|Q|}\int_Q |f - f_Q|\, dx \leq 2^{2n}$

Now suppose $\lambda > 4^n$. Then we observe:

$$\begin{aligned}
\{x \in Q : |f(x) - f_Q| > \lambda\} &\leq \sum_i |\{x \in Q_i : |f(x) - f_Q| > \lambda\}| \\
&\leq \sum_i |\{x \in Q_i : |f(x) - f_Q| > (\lambda - 4^n)\}| \frac{|Q_i|}{|Q_i|} \\
&\leq F(\lambda - 4^n) \sum_i |Q_i| \\
&\leq F(\lambda - 4^n) \frac{|Q|\,\|f\|_{BMO}}{2^n}
\end{aligned}$$

Therefore, we have found that:

$$|\{x \in Q : |f(x) - f_Q| \geq \lambda\}| \leq |Q| \frac{\|f\|_{BMO}}{2^n} F(\lambda - 4^n)$$

This holds for all $Q \subset \mathbb{R}^n$. Applying the definition of $F(\lambda)$, then we see,

$$|\{x \in Q : |f(x) - f_Q| \geq \lambda\}| \leq |Q|F(\lambda)$$

Therefore,

$$\begin{aligned} F(\lambda) &\leq \frac{\|f\|_{BMO}}{2^n} F(\lambda - 4^n) \\ &\leq \left(\frac{\|f\|_{BMO}}{2^n}\right)^2 F(\lambda - 24^n) && \text{Iterating Again} \\ &\leq \dots \end{aligned}$$

Let α be the largest integer such that $\lambda > \alpha 4^n \implies \alpha \geq \frac{\lambda}{4^n} - 1$. Therefore, applying the same iteration process α times :

$$F(\lambda) \leq \left(\frac{\|f\|_{BMO}}{2^n}\right)^\alpha F(\lambda - \alpha 4^n) = \exp\left(\alpha \ln(\|f\|_{BMO}) - \ln 2^n\right)$$

Case: Now assume that $\|f\|_{BMO} = 1$. Then

$$F(\lambda) = \exp(-\alpha \ln 2^n) \leq \exp(-(\frac{\lambda}{4^n} - 1)\ln 2^n) = 2^n \exp(\frac{-\ln 2^n}{4^n}\lambda)$$

Therefore, no matter the choice of λ

$$|\{x \in Q : |f - f_Q| > \lambda\}| \leq |Q| 2^n \exp(\frac{-\ln 2^n}{4^n}\lambda)$$

Case: Now, assume that $\|f\|_{BMO} \neq 0$. Then we can normalize our function by:

$$\bar{f} = \frac{f}{\|f\|_{BMO}}$$

Therefore,

$$|\{x \in Q : |f - f_Q| > \lambda\}| = |\{x \in Q : |\bar{f} - \bar{f}_Q| > \frac{\lambda}{\|f\|_{BMO}}\}| \leq |Q| 2^n \exp(\frac{-\ln 2^n}{4^n}\frac{\lambda}{\|f\|_{BMO}})$$

■

21.2 Dual Space of $H^s(\mathbb{R}^n)$

Recall we define the space $H^s(\mathbb{R}^n)$ with $s > 0$ by

$$H^s(\mathbb{R}^n) := \{u \in \mathcal{S}'(\mathbb{R}^n) : (1 + |\xi|^2)^{s/2}\hat{u} \in L^2\}$$

equipped with function: Recall $H^s(\mathbb{R}^n)$ is a Hilbert space, equipped with the inner product:

$$\langle u, v\rangle_{H^s} := \int (1 + |\xi|^2)^{\frac{s}{2}}\hat{u}(\xi)(1 + |\xi|^2)^{\frac{s}{2}}\hat{v}(\xi)\, d\xi = \langle (1 + |\xi|^2)^{\frac{s}{2}}\hat{u}, (1 + |\xi|^2)^{\frac{s}{2}}\hat{v}\rangle_{L^2}$$

So how can we go about finding the dual of $H^s(\mathbb{R}^n)$. Let $u \in (H^s(\mathbb{R}^n))'$. Then there exists $w \in H^s(\mathbb{R}^n)$ such that

$$\langle u, v\rangle = \langle w, v\rangle_{H^s(\mathbb{R}^n)}$$

But this implies that

$$\|u\|_{(H^s(\mathbb{R}^n))'} = \|w\|_{H^s(\mathbb{R}^n)}$$

Claim: $\|w\|_{H^s} = \|u\|_{H^{-s}}$
Because $u \in (H^s(\mathbb{R}^n))'$ and the schwartz class $\mathcal{S}(\mathbb{R}^n) \subset H^s(\mathbb{R}^n)$ is dense. Then we can simply consider $u \in \mathcal{S}'(\mathbb{R}^n)$, $\phi \in \mathcal{S}(\mathbb{R}^n)$ and therefore observe:

$$\begin{aligned}
\langle (1+|\xi|^2)^{\frac{-s}{2}}\hat{u}, \phi\rangle &= \langle u, (1+\widehat{|\xi|^2)^{\frac{-s}{2}}}\phi\rangle \\
&= \langle w, (1+\widehat{|\xi|^2)^{\frac{-s}{2}}}\phi\rangle \\
&= \int (1+|\xi|^2)^{\frac{s}{2}}\hat{w}(\xi)(1+|\xi|^2)^{\frac{s}{2}}\overline{(1+\widehat{\widehat{|\xi|^2)^{\frac{-s}{2}}}}\phi d\xi} \\
&= \int (1+|\xi|^2)^{\frac{s}{2}}\hat{w}(-\xi)\phi(\xi)d\xi \\
&= \langle (1+|\tilde{\xi}|^2)^{\frac{s}{2}}\hat{w}, \phi\rangle
\end{aligned}$$

where the tilde indicates reflection. Therefore,

$$\|u\|_{H^{-s}(\mathbb{R}^n)} = \left\|(1+|\xi|^2)^{\frac{-s}{2}}\hat{u}\right\|_{L^2} = \sup_{\|\phi\|=1}\langle (1+|\xi|^2)^{\frac{-s}{2}}\hat{u}, \phi\rangle = \left\|(1+|\xi|^2)^{\frac{s}{2}}\hat{w}\right\|_{L^2} = \|w\|_{H^s(\mathbb{R}^n)}$$

Therefore, we can conclude

$$(H^s(\mathbb{R}^n))' = H^{-s}(\mathbb{R}^n)$$

21.3 Hilbert Transform

The easiest way to define the Hilbert transform on a function is to define it's Fourier transform:

$$\widehat{Hf}(\xi) = -i \cdot sign(\xi)\hat{f}(\xi)$$

Geometrically this equates to a complex rotation of 90 degrees, since

$$i \cdot sign(\xi) = e^{-\frac{\pi}{2}\xi}$$

How do we then write $Hf(x) =?$. Firstly, suppose $f : \mathbb{C} \to \mathbb{C}$ an analytic function in $Im[z] > 0$. Define the functions on the boundary $u, v : \mathbb{R} \to \mathbb{R}$:

$$u(x) = Real[f(x + 0 \cdot i)] \quad v(x) = Im[f(x + 0 \cdot i)]$$

Then the critical conclusion is that:

$$v(x) = (Hu)(x) + C$$

with some control constant. Turning our attention to tempered distributions, we can write:

$$\widehat{Hf}(\xi) = -isign(\xi)\hat{f}(\xi) = \hat{h}(\xi)\hat{f}(\xi)$$

And therefore leveraging convolution, we see:

$$Hf(x) = (h * f)(x)$$

Therefore, to find h, we use the inversion formula:

$$h(\xi) - \int_{-\infty}^{\infty} isign(\xi)e^{2\pi i x\xi}d\xi = -i\widehat{sign}(-x) = \frac{-i}{\pi i(-x)} = \frac{1}{\pi x}$$

Therefore,

$$(Hf)(x) = \frac{1}{\pi x} * f$$

So if we were to calculate the principle value, we would see:

$$\begin{aligned} p.v.\frac{1}{\pi x} * f &= \frac{1}{\pi}\lim_{\epsilon\downarrow 0}\int_{|x-y|>\epsilon}\frac{f(y)}{x-y}\,dy \\ &= \frac{1}{\pi}\lim_{\epsilon\downarrow 0}\left[\int_{x-y>\epsilon}\frac{f(y)}{x-y}\,dy + \int_{x-y<-\epsilon}\frac{f(y)}{x-y}\,dy\right] \end{aligned}$$

Observation 21.4. *Notice, if f is even, $Hf = 0$*

Example 21.5. *Let $f = \mathcal{X}_{[a,b]}$ with $a < b$. Then*

$$Hf(x) = \frac{1}{\pi}\lim_{\epsilon\downarrow 0}\int_{|x-y|>\epsilon}\frac{\mathcal{X}_{[a,b]}}{x-y}\,dy = \frac{1}{\pi}\ln\left|\frac{x-a}{x-b}\right|$$

Observation 21.6. *The Hilbert Transform doesn't preserve functions in the space of L^∞. Instead, the Hilbert Transform takes the above function into the larger BMO space!*

Remark 21.7. *$f \in BMO \iff f = Hf_1 + f_2$ where $f_1, f_2 \in L^\infty(\mathbb{R})$*

Theorem 21.8. *Suppose $f \in \mathcal{S}(\mathbb{R}^n)$ and $f \geq 0$, then*

1. $|\{x \in \mathbb{R} : |Hf| > \lambda\}| \leq \frac{C\|f\|_{L^1}}{\lambda}$
2. $1 < p < \infty \implies \|Hf\|_{L^p} \leq C\,\|f\|_{L^p}$
3. $\|Hf\|_{BMO} \leq C\,\|f\|_{L^\infty}$

Proof. (1) $\implies$ (2)
We can use Marciekiwics Interpolation Theorem with Duality.
To prove for $1 < p \leq 2$, observe:

$$\|Hf\|_{L^2} = \left\|\widehat{Hf}\right\|_{L^2} = \left\|i\cdot sign(\xi)\hat{f}(\xi)\right\|_{L^2} = \left\|\hat{f}\right\|_{L^2} = \|f\|_{L^2}$$

On the other hand, we know $H : L^2 \to L^2$ and $H : L^1 \to L^{1,weak}$ are bounded, and therefore H is bounded for all $1 < p \leq 2$. Now to extend this, we show that the adjoint of H is $-H$:

$$\begin{aligned} \langle Hf, g\rangle_{L^2} &= \int Hf\cdot\overline{g} \\ &= \langle\widehat{Hf}, \hat{g}\rangle && \text{Plancheral} \\ &= \int i\cdot sign(\xi)\hat{f}(\xi)\overline{\hat{g}(\xi)}d\xi \\ &= \int \hat{f}(\xi)\overline{-i\cdot sign(\xi)\hat{g}(\xi)}d\xi \\ &= \langle\hat{f}, -\widehat{Hg}\rangle \\ &= \langle f, -Hg\rangle \end{aligned}$$

Now, to prove for $2 \leq p < \infty$, we recall the Hölder complement

$$\frac{1}{p} + \frac{1}{q} = 1 \implies 1 < q \leq 2$$

Therefore,

$$\begin{aligned}\|Hf\|_{L^p} &= \sup_{\|g\|_q=1} \int Hf \cdot g \\ &= \sup_{\|g\|_q=1} \int f \cdot (-Hg) \\ &\leq \sup_{\|g\|_q=1} \|f\|_{L^p} \|Hg\|_{L^q} \\ &\leq \|f\|_{L^p}\end{aligned}$$

Therefore, by duality, we see:

$$\|Hg\|_{L^q} \leq 1$$

We will return for the proof of (1) later. ∎

21.4 Riesz Transform

Generalizing the Hilbert Transform on $\mathbb{R}^n$:

$$\widehat{R_i f}(\xi) := i\frac{\xi_i}{|\xi|}$$

with $1 \leq i \leq n$ on a function $f : \mathbb{R}^n \to \mathbb{R}$. Returning this from the frequency domain, the transform is written

$$R_i f(x) = \frac{1}{\pi} \lim_{\epsilon \downarrow 0} \int_{\mathbb{R}^n} \frac{y_i}{|y|^{n+1}} f(x-y)dy$$

Further, $R_i : L^p \to L^p$ with $1 < p < \infty$ is bounded!
This operator has a great application. Consider the defining PDE:

$$\begin{cases}\Delta u = f\end{cases}$$

Then upon taking the Fourier Transform, we see:

$$-2\pi \sum_{i=1}^n \xi_i^2 \hat{u} = \hat{f}$$

$$\implies \hat{u} = \frac{\hat{f}}{-(2\pi)^2|\xi|^2}$$

On the other hand,

$$\widehat{\partial_{x_i}\partial_{x_j}} u = -(2\pi)^2 \xi_i \xi_j \hat{u} = \widehat{R_i R_j} f$$

Therefore the Riesz operator is a differential operator. Therefore, we see:

$$\partial_{x_i} \partial x_j u = R_i R_j f$$

Further,

$$\|\partial_{x_i} \partial x_j u\|_{L^p} \leq \|f\|_{L^p}$$

21.5 Calderon-Zygmund Operator

Theorem 21.9. *Suppose K is a tempered distribution that satisfies the properties:*

1. $|\hat{K}(\xi)| \leq C < \infty$
2. $\int_{|x|>2|y|} |K(x-y) - K(x)|dx \leq C$ *(Hörmauder's Condition)*

Then

1. $K * f$ *is bounded from* $L^1 \to L^{1,weak}$
2. $\|K * f\|_{L^p} \leq C \|f\|_{L^p}$ *is bounded from* $L^p \to L^p$ *with* $1 < p < \infty$.

Proof. So conclusion (1) implies (2) via Marcinkiewicz.
To prove that $K * f : L^1 \to L^{1,weak}$ is bounded, we will use the Calderon-Zygmund decomposition theorem. That is, we need to prove,

$$|\{x : |(K * f)(x)| > \lambda\}| \leq \frac{C \|f\|_{L^1}}{\lambda}$$

for all $\lambda > 0$. So we recall that there exists cubes $\{Q_i\}$ that are pairwise disjoint such that the following hold:

1. $|f(x)| \leq \lambda$ for almost every $x \in (\bigcup_i Q)^c$
2. $|\bigcup_i Q_i| \leq \frac{C\|f\|_{L^1}}{\lambda}$
3. $\lambda < \frac{1}{|Q_i|} \int_{Q_i} f \leq 2^n \lambda$

Now fix λ by Calderan-Zygmund Decomposition. Then we can decompose

$$f = g + b$$

where

1. $g \in L^\infty$ and $\|g\|_{L^\infty} \leq 2^n \lambda$
2. $supp(b) \subset \bigcup_i Q_i$
3. $\int_{Q_i} b = 0$ for each Q_i.

More specifically, we may let

$$g(x) = \begin{cases} f(x) & x \in (\bigcup_i Q_i)^c \\ \frac{1}{|Q_i|} \int_{Q_i} f & x \in \bigcup_i Q_i \end{cases}$$

And then simply allow

$$b(x) = f(x) - g(x) = \begin{cases} 0 & x \in (\bigcup_i Q_i)^c \\ f(x) - \frac{1}{|Q_i|} \int_{Q_i} f & x \in \bigcup_i Q_i \end{cases}$$

Therefore, we see since $K * f = K * g + K * b$

$$\implies |\{x : |(K * f)(x)| > \lambda\}| \leq \underbrace{|\{x : |(K * g)(x)| > \frac{\lambda}{2}\}|}_{I} + \underbrace{|\{x : |(K * b)(x)| > \frac{\lambda}{2}\}|}_{II}$$

$$
\begin{aligned}
I &= |\{x : |(K*g)(x)| > \frac{\lambda}{2}\}| \\
&= |\{x : |(K*g)(x)|^2 > \frac{\lambda^2}{2^2}\}| \\
&\leq \frac{4}{\lambda^2}\|K*g\|_{L^2}^2 && \text{Chebyshev} \\
&= \frac{4}{\lambda^2}\left\|\hat{k}\hat{g}\right\|_{L^2}^2 \\
&= \frac{4C^2\|g\|_{L^2}^2}{\lambda^2} \\
&\leq C\frac{\|g\|_{L^\infty}\|g\|_{L^1}}{\lambda^2} \\
&\leq \frac{C2^n\lambda\|f\|_{L^1}}{\lambda^2} \\
&\leq \frac{C'\|f\|_{L^1}}{\lambda}
\end{aligned}
$$

$$
\begin{aligned}
II &\leq |\{x \in \mathbb{R}^n : |K*b| > \frac{\lambda}{2}\}| \\
&= |\{x \in \bigcup_i Q_i : |K*b| > \frac{\lambda}{2}\}| + |\{x \in \mathbb{R}^n \setminus \bigcup_i Q_i : |K*b| > \frac{\lambda}{2}\}|
\end{aligned}
$$

Now, let Q_i^* represent concentric dilations of the cube Q_i by a factor of $2\sqrt{n}$.. Notice, if $y \in Q_i$, and $x \in (Q_i^*)^c$, then $|x - c_1| \geq 2|y - c_1|$. So we see:

$$
\begin{aligned}
|\{x \in \mathbb{R}^n \setminus \bigcup_i Q_i : |K*b| > \frac{\lambda}{2}\}| &\leq \frac{2}{\lambda}\int_{x \notin \bigcup_i Q_i} |K*b|dx && \text{Che} \\
&\leq \frac{2}{\lambda}\sum_i \int_{x \notin \bigcup_i Q_i} |K*b_i|dx && b = \sum_i b\mathcal{X}_{Q_i} \\
&= \frac{2}{\lambda}\sum_i \int_{x \notin \bigcup_i Q_i} \left|\int_{\mathbb{R}^n} K(x-y)b_i(y)\ dy\right|\ dx \\
&\leq \frac{2}{\lambda}\int_{x \notin Q_i^*}\int_Q (K(x-y)b_i(y) - K(x-c_i)b_i(y))dydx \\
&= \frac{2}{\lambda}\int_{y \in Q_i} |b_i(y)| \int_{|x-c_i|>2|y-c_i|} |K(x-c_i-(y-c_i)) - K(x-c_i)|dx\ dy \\
&\leq \sum_i \frac{2c}{\lambda}\|b_i\|_{L^1(Q_i)} \\
&= \frac{2c}{\lambda}\|b\|_{L^1(\mathbb{R}^n)}
\end{aligned}
$$

Therefore, $\|b\|_{L^1(\mathbb{R}^n)} \leq 2\|f\|_{L^1(\mathbb{R}^n)}$. ■

Theorem 21.10. *If* $|\nabla K(x)| \leq \frac{C}{|x|^{n+1}}$*, then* $\int_{|x|>2|y|} |K(x-y) - K(x)|dx \leq C$ *holds.*

Proof. Case: $|\nabla K(x)| \leq \frac{C}{|x|^{n+1}} \implies \int_{|x|>2|y|} |K(x-y) - K(x)|dx \leq C$ Therefore,

$$
\int_{|x|>2|y|} |K(x-y) - K(x)|dx = \int_{|x|>2|y|} |\nabla K(\xi) \cdot (-y)|dx
$$

Now let $\xi = tx + (1-t)y$ for $0 \le t \le 1$. Then

$$\frac{|x|}{2} \le |\xi|$$

But this implies:

$$|\nabla K(\xi)| \le \frac{C}{|\xi|^{n+1}} \le \frac{C'}{|x^{n+1}|}$$

Therefore, we return to the chain:

$$\begin{aligned}
\int_{|x|>2|y|} |K(x-y) - K(x)| dx &= \int_{|x|>2|y|} |\nabla K(\xi) \cdot (-y)| dx \\
&\le \int_{|x|>2|y|} \frac{C'}{|x|^{n+1}} |y| dx \\
&\le |y| \int_{|x|>2|y|} \frac{C'}{|x|^{n+1}} x^{n+1} dx \\
&= |y| \int_{|x|>2|y|} \frac{C'}{|x|^2} dx \\
&\le \frac{C'|y|}{|y|} = C'
\end{aligned}$$

■

21.6 Last Lecture

This course is effectively the study of function spaces. The first spaces we studied were the L^p spaces. Thinking about this space, we see:

- $\|f\|_{L^p}^p = \int_{\mathbb{R}^n} |f|^p < \infty$
- $\|f\|_{L^p} = \sup_{\|g\|_{L^q}} \int fg$ (Riesz-Representation which leads to weak convergence)
- $\|f\|_{L^p} = p \int t^{p-1} \mu\{f(x) > t\}\ dt$

The second space we considered were Sobolev Spaces $W^{k,p}(\mathbb{R}^n)$. When $k = 1$, we have the following relationships:

- $1 \le p \le n \implies$ the space is embeds the p^* space

$$\|f\|_{L^{p^*}} \le \|\nabla f\|_{L^p}$$

- $p > n \implies W^{1,p}(\mathbb{R}^n) \subset C^{0,1-\frac{n}{p}}$
- $p = n \implies W^{1,n}(\mathbb{R}^n) \subset$ BMO (Bounded Min Oscillation) Space

We then define the Hilbert Transform on the BMO space to be:

$$BMO = HilbertTransform(L^\infty) + L^\infty$$

Restriction and Extension Theorems in Sobolev Spaces:

- $f \in C^m(\mathbb{R}^n) \implies f|_\Omega \in C^m(\Omega)$
- $f \in W^{1,p}(\mathbb{R}^n) \implies f|_{\mathbb{R}^{n-1}} \notin W^{1,p}(\mathbb{R}^{n-1})$ by the Trace Theorem
- Use Fourier Transform to define the fractional Sobolev Space $H^s = W^{s,2}(\mathbb{R}^n)$

- Taking the space of tempered distributions $(H^s)^* = H^{-s}$.
- $H^s(\mathbb{R}^n)|_{\mathbb{R}^{n-1}} = H^{s-1/2}(\mathbb{R}^{n-1})$

Recall the Fourier transform was the starting point to solving partial differential equations. Given a fourier transform, what are the conditions that which we can recover the original function?

- The space of Schwartz Functions $\mathcal{S}(\mathbb{R}^n)$ allows for interchangeability
- $f, \hat{f} \in L^1 \implies f$ can be reconvered from $\hat{f}$ a.e.

If we want to analyze a given function in L^1, we can decompose it into nice pieces to understand it better. By CZD:

$$f = g + b$$

where $g \in L^\infty$ and there exist cubes $\{Q_i\}$ such that $f_{Q_i} b = 0$.

If you want to study the relationships between L^p and L^q, then we can use the Real Interpolation Theorem (Marvinkiewics). If we use the Complex Integral form, then the Riesz Thorin Interpolation can serve us well. These tools helped us from the bounds of the Fourier Transform and Convolution.

We can then use the Fourier Transform to solve for the solution of the Dirichlet problem on Ω within the space $W^{1,2}(\Omega)$!