MAT 425 - Integration Theory + Hilbert Spaces

OH: 4-5pm Mon, Fine 707

PSETS: Fridays 11:59pm

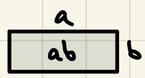
Lecture 1/30 - First day yippee!

The heart of many analysis questions is the following: "what even is area?"
We answer this w/ the Lebesgue Measure.

§1: Lebesgue Measure

§1.1: Preliminaries

It is reasonable to say that the rectangle $[a_1,b_1]\times[a_2,b_2]\times\dots\times[a_n,b_n]\subset\mathbb{R}^n$ has area $\prod_{i=1}^{n}(b_i-a_i)$.



This will be our starting point.

Def: A (closed) rectangle $R\subset\mathbb{R}^n$ is a set of the form

$$R=[a_1,b_1]\times\dots\times[a_n,b_n] \quad \text{with} \quad a_i<b_i \ \forall i$$

The volume is then
$$|R|=\prod_{i=1}^{n}(b_i-a_i)$$

The interior of R is
$$\text{int}(R)=(a_1,b_1)\times\dots\times(a_n,b_n)$$

Def: A collection of (closed) rectangles $\{R_\alpha\}_{\alpha\in I}$ is almost disjoint if $\forall\alpha,\beta$
$$\text{int}(R_\alpha)\cap\text{int}(R_\beta)=\emptyset$$

Note: I must be countable because each interior contains a rational point, if the interiors are nonempty.

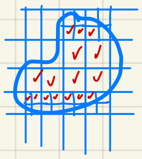
From these definitions, we can prove:

Lemma 1.1: If R is a rectangle which is an almost disjoint union of finitely many other rectangles $R=\bigcup_{k=1}^{N}R_k$, then $|R|=\sum_{k=1}^{N}|R_k|$

Lemma 1.2: If $R,(R_k)_{k=1}^{N}$ are rectangles with $R\subseteq\bigcup_{k=1}^{N}R_k$, then $|R|\le\sum_{k=1}^{N}|R_k|$

From here, we can extend to more general sets via

Theorem 1.4: Every *open* set $U \subseteq \mathbb{R}^n$ can be written as a (countable) union of almost disjoint cubes $U = \bigcup_{i=1}^{\infty} R_i$, where $\{R_i\}_{i=1}^{\infty}$ are almost disjoint.

Proof sketch:  Start w/ $\mathbb{Z}^n$ lattice grid. Save the cubes contained in U, and bisect all the cubes partially contained in U. Iterate this. U open $\Rightarrow \forall x \in U$, x will lie in some small enough cube, which is entirely contained in U.

It is reasonable to hope to define the $\mathrm{Vol}(U)$ as the sum of these areas. We have to check that all the different rectangularizations yield the same volume.

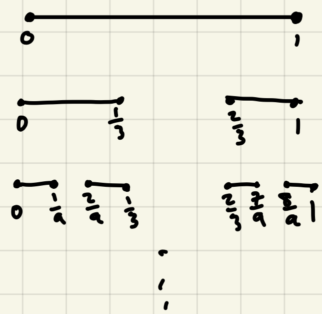
Key example (Cantor Set): Remove the middle third open intervals to get

$C_0 = [0,1]$

$C_1 = [0,\frac{1}{3}] \cup [\frac{2}{3},1]$

$C_2 = [0,\frac{1}{9}] \cup [\frac{2}{9},\frac{1}{3}] \cup [\frac{2}{3},\frac{7}{9}] \cup [\frac{8}{9},1]$

$\vdots$



The Cantor Set is defined by $C = \bigcap_{i=0}^{\infty} C_i$, and enjoys the following properties

1. C closed
2. C bounded

$\Rightarrow C$ is compact

3. C is totally disconnected (only connected subsets are singletons)
4. C is uncountable

From a cardinality perspective, C is huge. From an area perspective, the area of each C_n is $\left(\frac{2}{3}\right)^n$ (2^n intervals of length $\frac{1}{3^n}$), and so the area of C should be 0.

Let us make this precise.

§1.2 - Exterior Measure

Idea: get a first attempt of volume by taking coverings by cubes and taking an infimum.

Def: For a subset $E \subseteq \mathbb{R}^n$, define the exterior measure of E by

$$m_*(E) = \inf \sum_{j=1}^{\infty} |Q_j|$$

where the inf is over all countable coverings of E by closed cubes $E \subseteq \bigcup_{i=1}^{\infty} Q_i$

Properties of m_*: • $0 \le m_*(E) \le +\infty$
• inf must be over <u>countable</u> coverings, not just finite
• one can also work with rectangles to get same results

Lecture 2/1 -

<u>m_* examples</u>

a) $m_*(\{\text{point}\})=0$

b) If C is a closed cube, then $m_*(C)=|C|$

<u>Proof:</u> $C \subseteq C \Rightarrow m_*(C) \le |C|$

For the other direction, $\forall \varepsilon > 0$ we can take a cover $C \subseteq \bigcup_{i=1}^{\infty} Q_i$ with $\sum_{i=1}^{\infty} |Q_i| \le m_*(C) + \varepsilon$. For each Q_i we can take an open cube $S_i \supseteq Q_i$ with $|S_i| \le |Q_i| + \frac{\varepsilon}{2^i}$ $\Rightarrow$ $C \subseteq \bigcup_{i=1}^{\infty} S_i$ ← open cover of compact set ⇒ finite subcover

So, there is some finite index set I s.t.

$$C \subseteq \bigcup_{i \in I} S_i \Rightarrow |C| \le \sum_{i\in I} |S_i| \le \sum_{i=1}^{\infty} |S_i| \le \sum_{i=1}^{\infty} |Q_i| + \frac{\varepsilon}{2^i} = \varepsilon + \sum_{i=1}^{\infty} |Q_i| \le 2\varepsilon + m_*(C)$$

□

Common technique: take a near-infimum cover, modify it a bit and roll

c) If C is the Cantor set, then $m_*(C)=0$.

<u>Note:</u> at the moment, the exterior measure isn't additive under the countable union of disjoint subsets.

<u>Prop. 1.</u>: (Properties of m_*)

① (Monotonicity) $E_1 \subseteq E_2 \Rightarrow m_*(E_1) \le m_*(E_2)$

② (Countable subadditivity) $E = \bigcup_{i=1}^{\infty} E_i \Rightarrow m_*(E) \le \sum_{i=1}^{\infty} m_*(E_i)$

③ $m_*(E) = \inf_{\substack{U \supseteq E \\ U \text{ open}}} m_*(U)$

④ $\text{dist}(E_1,E_2)>0 \Rightarrow m_*(E_1 \cup E_2) = m_*(E_1) + m_*(E_2)$

⑤ if $E = \bigcup_{i=1}^{\infty} Q_i$ where $\{Q_i\}_{i=1}^{\infty}$ are almost disjoint cubes, then
$$m_*(E) = \sum_{i=1}^{\infty} m_*(Q_i)$$

<u>Remark:</u> ⑤ tells us that the definition of volume of open sets from last time is well-defined.

Proof: ① Any cover of E_2 is a cover of E_1, and so the inf over covers of E_1 will be smaller.

② WOLOG assume $\sum_i m_*(E_i)$ is finite. Fix $\varepsilon>0$. Then, $\forall i$ choose a cover $E_i \subseteq \bigcup_j Q_{ij}$ of closed cubes Q_{ij} with $\sum_j |Q_{ij}| \leq m_*(E_i) + \frac{\varepsilon}{2^i}$

again w/ the take a bigger cover trick!

$$\Rightarrow E \subseteq \bigcup_{i,j} Q_{ij} \Rightarrow m_*(E) \leq \sum_{i,j} |Q_{ij}| \leq \varepsilon + \sum_i m_*(E_i)$$

Taking $\varepsilon \to 0$ yields the result.

③ Clearly, $E \subseteq U \Rightarrow m_*(E) \leq m_*(U)$ $\forall U$, by ①. So, $m_*(E) \leq \inf_{\substack{U \supseteq E \\ U \text{ open}}} m_*(U)$

For the other direction, fix $\varepsilon>0$. WOLOG suppose $m_*(E)$ finite. We can find $\{Q_i\}_{i=1}^{\infty}$ of closed cubes s.t. $E \subseteq \bigcup_{i=1}^{\infty} Q_i$ and $\sum_{i=1}^{\infty} |Q_i| \leq m_*(E)+\varepsilon$

For each i take an open cube $S_i \supseteq Q_i$ with $|S_i| \leq |Q_i| + \frac{\varepsilon}{2^i}$. Then,

$$E \subseteq \underbrace{\bigcup_{i=1}^{\infty} S_i}_{\text{open}} =: U, \text{ and by ② } m_*(U) \leq \sum_{i=1}^{\infty} m_*(S_i) = \sum_{i=1}^{\infty} |S_i| \leq \varepsilon + \sum_{i=1}^{\infty} |Q_i| \leq m_*(E) + 2\varepsilon$$

Taking $\varepsilon \to 0$ yields $\inf_{\substack{U \supseteq E \\ U \text{ open}}} m_*(U) \leq m_*(E)$

④ Subadditivity yields $m_*(E) \leq m_*(E_1) + m_*(E_2)$. For the other direction, take $\{Q_i\}_{i=1}^{\infty}$ to be closed cubes with $E \subseteq \bigcup_{i=1}^{\infty} Q_i$ and $\sum_{i=1}^{\infty} |Q_i| \leq m_*(E)+\varepsilon$ for some fixed $\varepsilon>0$.

If $\delta = \frac{1}{2}\operatorname{dist}(E_1,E_2)$, by bisecting the Q_i's iteratively we can WOLOG assume that all the Q_i have side lengths $<\delta$. So, each Q_i can intersect at most one of E_1 or E_2. Let I_1 be $\{i \mid Q_i \cap E_1 \neq \emptyset\}$ and $I_2 = \{i \mid Q_i \cap E_2 \neq \emptyset\} \Rightarrow I_1 \cap I_2 = \emptyset$.

Then, $E_j \subseteq \bigcup_{i \in I_j} Q_i$ for $j \in \{1,2\}$. So, $m_*(E_j) \leq \sum_{i \in I_j} |Q_j|$ by def. of exterior measure.

So, $m_*(E_1) + m_*(E_2) \leq \sum_{i \in I_1} |Q_i| + \sum_{i \in I_2} |Q_i| \underset{\substack{\text{because} \\ I_1, I_2 \text{ disjoint}}}{\leq} \sum_{i=1}^{\infty} |Q_i| \leq m_*(E)+\varepsilon$. Take $\varepsilon \to 0$.

⑤ ($\leq$) comes once again from countable subadditivity. For the other direction, fix $\varepsilon>0$. For each Q_i, take a closed cube $S_i \subseteq Q_i$ s.t. $|S_i| \geq |Q_i| - \frac{\varepsilon}{2^i}$. Since we shrunk to get the S_i's, each pair (S_i, S_j) has distance >0. We can then apply ④ inductively to get that any finite union has $m_*\left(\bigcup_{i=1}^{N} S_i\right) = \sum_{i=1}^{N} |S_i| \geq \sum_{i=1}^{N} |Q_i| - \frac{\varepsilon}{2^i}$, $N \in \mathbb{N}$

Since $\bigcup_{i=1}^{N} S_i \subseteq \bigcup_{i=1}^{\infty} Q_i = E$, monotonicity ① gives $m_*(E) \geq m_*\left(\bigcup_{i=1}^{N} S_i\right) \geq \sum_{i=1}^{N} |Q_i| - \varepsilon$

Taking $N \to \infty$, $\varepsilon \to 0$ gives $m_*(E) \geq \sum_{i=1}^{\infty} |Q_i|$ yields the last result. $\square$

§1.3 - Measurable Sets + Lebesgue Measure

Currently, m_* is <u>not countably additive</u> on disjoint sets. (see the Vitali sets)

Def: A set $E \subseteq \mathbb{R}^n$ is (Lebesgue) measurable if $\forall \varepsilon > 0$, $\exists$ an open set U with

$$E \subseteq U \quad \text{and} \quad m_*(U \setminus E) \leq \varepsilon$$

Def: If $E \subseteq \mathbb{R}^n$ is measure, its (Lebesgue) measure is $m(E) := m_*(E)$

Remarks:
- Prop 1's properties are inherited by $m(\cdot)$
- U open $\Rightarrow$ U measurable by definition
- $m_*(E) \Rightarrow E$ measurable by property ③

Prop. 2 ("Closure" Properties)

① A countable union of measurable sets is measurable

② Closed sets are measurable

③ Complement of a measurable set is measurable

④ A countable intersection of measurable sets is measurable

Proof:

① Suppose $\{E_i\}_{i=1}^{\infty}$ are measurable. Fix $\varepsilon > 0$. $\forall i$, $\exists$ an open set $U_i \supseteq E_i$ with $m_*(U_i \setminus E_i) \leq \frac{\varepsilon}{2^i}$. Set $U = \bigcup_{i=1}^{\infty} U_i$; then, $U \supseteq \bigcup_{i=1}^{\infty} E_i$ is open.

Note that $U \setminus \left(\bigcup_{i=1}^{\infty} E_i\right) \subseteq \bigcup_{i=1}^{\infty} (U_i \setminus E_i) \xRightarrow[\text{monotonicity}]{} m\left(U \setminus \bigcup_{i=1}^{\infty} E_i\right) \leq m_*\left(\bigcup_{i=1}^{\infty} (U_i \setminus E_i)\right) \leq \sum_{i=1}^{\infty} m_*(U_i \setminus E_i) \leq \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} = \varepsilon.$

Lecture 2/6 starts here

② Let C be closed. Then, $C = \bigcap_{i=1}^{\infty} \underbrace{(C \cap \overline{B}_i(0))}_{\text{closed + bounded}}$. Using (1), it suffices to prove that a compact $K \subseteq \mathbb{R}^n$ is measurable. Since K compact, $m_*(K) < \infty$.
Fix $\varepsilon > 0$. By Prop. 1(c), we can find $U \supseteq K$ s.t. U open and $m_*(U) \leq m_*(K) + \varepsilon$
We know $U \setminus K$ is open, and so by Thm. 1.4, $U \setminus K = \bigcup_{i=1}^{\infty} Q_i$ for $\{Q_i\}$ almost disjoint cubes.
Q_i is compact $\forall i$ and disjoint from K $\forall N \geq 1$, $\bigcup_{i=1}^{N} Q_i$ is compact and disjoint from K, which is also compact. So, $\text{dist}\left(\bigcup_{i=1}^{N} Q_i, K\right) > 0$. By Prop. 1, $m_*\left(\bigcup_{i=1}^{N} Q_i \cup K\right) = m_*\left(\bigcup_{i=1}^{N} Q_i\right) + m_*(K)$

$\Rightarrow m_*(U) \geq \sum_{i=1}^{N} m_*(Q_i) + m_*(K) \Rightarrow \sum_{i=1}^{N} m_*(Q_i) \leq \varepsilon.$

Letting $N \to \infty$, $\sum_{k=1}^{\infty} m_*(Q_i) \leq \varepsilon \Rightarrow m_*\left(\bigcup_{i=1}^{\infty} Q_i\right) = m_*(U \setminus K) \leq \varepsilon.$

(3) Let E be measurable. $\forall k \geq 1$, $\exists U_k \supseteq E$ open s.t. $m_*(U_k \setminus E) \leq \frac{1}{k}$.
So, $U_k^c \subseteq E^c$ is closed, and by (2) is measurable. By (1), $S = \bigcup_{k=1}^{\infty} U_k^c$ is measurable.

Also, $E^c \setminus S \subseteq U_k \setminus E$ $\forall k$ $\rightarrow m_*(E^c \setminus S) \leq m_*(U_k \setminus E) \leq \frac{1}{k}$ $\forall k \Rightarrow m_*(E^c \setminus S) = 0$
$\Rightarrow E^c \setminus S$ is measurable $\underset{(1)}{\Rightarrow} E^c = (E^c \setminus S) \cup S$ is measurable.

this technique shows that E^c is measurable by showing that it differs from a measurable set (S) by a set $(E^c \setminus S)$ of measure 0.

(4) As $\bigcap_{k=1}^{\infty} E_k = \left(\bigcup_{k=1}^{\infty} E_k^c\right)^c$, (1) and (3) gives (4).

$\square$

Remark: not true for uncountable union/intersection

We can **finally** prove what we want $m(\cdot)$ to have!

Theorem 3.2 (Additivity of Measure)

If $\{E_k\}_{k \in \mathbb{N}}$ are measurable, disjoint sets, then
$$m\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} m(E_k)$$

Proof: Subadditivity gives $(\leq)$. We wish to bound $(\geq)$.
Case 1: Suppose all E_k are bounded. E_k measurable $\Rightarrow E_k^c$ measurable.
Fix $\varepsilon > 0$. Choose open $U_k \supseteq E_k^c$ s.t. $m_*(U_k \setminus E_k^c) \leq \frac{\varepsilon}{2^k}$.
Then, $F_k := U_k^c \subseteq E_k$, F_k is closed, and $E_k \setminus F_k = U_k \setminus E_k^c$.
So, $m_*(E_k \setminus F_k) \leq \frac{\varepsilon}{2^k}$

Now, E_k bounded $\Rightarrow F_k$ compact. Also, E_k disjoint $\Rightarrow F_k$ disjoint.
This yields disjoint compact sets, which are a positive distance apart.
For any finite $N \in \mathbb{N}$, Prop. 1 yields $m_*\left(\bigcup_{k=1}^{N} F_k\right) = \sum_{k=1}^{N} m_*(F_k)$.

Also, $\bigcup_{i=1}^{\infty} E_i \supseteq \bigcup_{i=1}^{N} F_i \Rightarrow m_*\left(\bigcup_{i=1}^{\infty} E_i\right) \geq \sum_{k=1}^{N} m_*(F_k)$

Subadditivity yields $m_*\left(\bigcup_{i=1}^{\infty} E_i\right) \geq \sum_{i=1}^{N} \left(m_*(E_i) - \frac{\varepsilon}{2^i}\right)$

Taking $N \to \infty$, $\varepsilon \to 0$, we are done with the bounded case.

Case 2: general case
Set S_1 to be the closed unit ball, $S_j = \overline{B_j(0)} \setminus \overline{B_{j-1}(0)}$
Then, $\{S_j\}$ are disjoint and $\mathbb{R}^n = \bigcup_{j=1}^{\infty} S_j$.

S_j are measurable, so set $E_{jk} := E_j \cap S_k$
Then, E_{jk} are disjoint, bounded measurable sets with $\bigcup_{i=1}^{\infty} E_i = \bigcup_{j,k} E_{jk}$

case 1 again

By Case 1, $m_*\left(\bigcup_{i=1}^{\infty} E_i\right) = m_*\left(\bigcup_{j,k} E_{jk}\right) = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} m_*(E_{jk}) = \sum_{j=1}^{\infty} m_*\left(\bigcup_{k=1}^{\infty} E_{jk}\right) = \sum_{j=1}^{\infty} m_*(E_j)$

$\square$

technique: general case comes from bounded case by exhausting $\mathbb{R}^n$ with bounded, disjoint thingies (S_j)

We now know that Lebesgue measure isn't stupid. Let us examine further properties.

Corollary 3.? (Further Properties of $m(\cdot)$):

Suppose $\{E_i\}_{i\in\mathbb{N}}$ are measurable.

(i) if $\{E_i\}_i$ is increasing $(E_i \subseteq E_{i+1}\ \forall i)$, then
$$m\left(\bigcup_{i=1}^{\infty} E_i\right) = \lim_{i\to\infty} m(E_i)$$

(ii) if $\{E_i\}_i$ is decreasing $(E_i \supseteq E_{i+1}\ \forall i)$ and $m(E_i)<\infty$ for some i,
$$m\left(\bigcap_{i=1}^{\infty} E_i\right) = \lim_{i\to\infty} m(E_i)$$

Remark: The condition $m(E_i)<\infty$ in (ii) is necessary and nontrivial. E.g. $E_i := (i,\infty)$. This phenomenon is like the measure "loses mass" at ∞, as the measure gets pushed toward ∞

Proof: (i) Set $G_1 := E_1$ and $G_i := E_i \setminus E_{i-1}$ $\forall i\geq 2$. Then, $\{G_i\}$ measurable and disjoint and $\bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} G_i \Rightarrow m\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} m(G_i) = \lim_{N\to\infty}\sum_{i=1}^{N} m(G_i)$
$$= \lim_{N\to\infty} m\left(\bigcup_{i=1}^{N} G_i\right) = \lim_{N\to\infty} m(E_N)$$

(ii) If $G_i = E_i \setminus E_{i+1}$ $\forall i\geq 1$, then $E_1 = \left(\bigcap_{i=1}^{\infty} E_i\right) \cup \left(\bigcup_{i=1}^{\infty} G_i\right)$

all disjoint

WOLOG, $m(E_1)<\infty \Rightarrow \cancel{m(E_1)} = m\left(\bigcap_{i=1}^{\infty} E_i\right) + \sum_{i=1}^{\infty} m(G_i)$
$$= m\left(\bigcap_{i=1}^{\infty} E_i\right) + \lim_{N\to\infty}\sum_{i=1}^{N}\left(m(E_i) - m(E_{i+1})\right)$$
$$= m\left(\bigcap_{i=1}^{\infty} E_i\right) + \lim_{N\to\infty}\left[\cancel{m(E_1)} - m(E_{N+1})\right]$$

can only cancel because $m(E_1)<\infty$

$$\Rightarrow m\left(\bigcap_{i=1}^{\infty} E_i\right) = \lim_{N\to\infty} m(E_n)$$

$\square$

Theorem 3.4 (More implications of measurability)

Suppose $E \subseteq \mathbb{R}^n$ is measurable. Then, $\forall \varepsilon > 0$,

(i) $\exists\, U \supseteq E$ open with $m_*(U \setminus E) \leq \varepsilon$

(ii) $\exists\, C \subseteq E$ closed with $m_*(E \setminus C) \leq \varepsilon$

(iii) If $m(E) < \infty$, then $\exists\, K \subseteq E$ compact with $m(E \setminus K) \leq \varepsilon$

(iv) If $m(E) < \infty$, then $\exists\, F = \bigcup_{j=1}^{N} Q_j$ finite union of closed cubes with $\underbrace{m(E \Delta F)}_{\text{symmetric diff. (XOR)}} \leq \varepsilon$

Proof: (i) is definition.

(ii) we saw before: follows from measurability of E^c and taking complements

(iii) As $E \cap \overline{B_R}(0)$ increases to E as $R \to \infty$, Corollary 3.? gives $\exists\, R > 0$ s.t. $m_*(\underbrace{E \setminus (E \setminus \overline{B_R}(0))}_{\text{measure of } E \text{ outside the ball}}) \leq \varepsilon$.

Applying (ii), we can find $K \subseteq E \setminus \overline{B_R}(0)$ closed (thus compact) with

$m_*(E \cap \overline{B_R}(0) \setminus K) \leq \varepsilon \Rightarrow m_*(E \setminus K) \leq 2\varepsilon$.

(iv) By defn. of $m_*(\cdot)$, $\exists\, E \subseteq \bigcup_{j=1}^{\infty} \underbrace{Q_j}_{\text{closed cubes}}$ with $\sum_{j=1}^{\infty} |Q_j| \leq m_*(E) + \varepsilon$

The fact that $\sum_{j=1}^{\infty} |Q_j|$ converges allows us to take large enough N that,

$\sum_{j=N+1}^{\infty} |Q_j| \leq \varepsilon$. Let $F = \bigcup_{j=1}^{N} Q_j \Rightarrow m(E \Delta F) = m(\underbrace{E \setminus F}_{\subseteq \bigcup_{j=N+1}^{\infty} Q_j}) + m(F \setminus E)$

$$\leq m\Big(\bigcup_{j=N+1}^{\infty} Q_j\Big) + m\Big(\bigcup_{j=1}^{\infty} Q_j \setminus E\Big)$$

$$\leq \sum_{j=N+1}^{\infty} |Q_j| + \sum_{j=1}^{\infty} |Q_j| - m(E) \leq 2\varepsilon$$

$\blacksquare$

Lecture 2/8.

Remark: The Lebesgue measure is translation invariant (by definition)

$$m(A+a) = m(A) \quad \forall a \in \mathbb{R}^n \text{ and } A \text{ measurable}$$

We may wish to study how complicated measurable sets are

Defn: A σ-algebra is a collection of sets that is closed under taking complements and countable unions (and thus countable intersections).

So, the measurable sets $\mathcal{M}$ form a σ-algebra.

Defn: The Borel σ-algebra $\mathcal{B}_{\mathbb{R}^n}$ is the smallest σ-algebra containing all open sets. Elements of $\mathcal{B}_{\mathbb{R}^n}$ are Borel sets.

Clearly $\mathcal{B}_{\mathbb{R}^n} \subseteq \mathcal{M}$. This inclusion is strict.

☆ Fact: $\mathcal{M}$ is the "completion" of $\mathcal{B}_{\mathbb{R}^n}$ by adding in subsets of Borel sets with measure 0.

Corollary 3.5:

The following are equivalent:

(i) $E \subseteq \mathbb{R}^n$ is measurable

countable intersection of open sets

(ii) E differs from a G_δ set by a set of measure 0

(iii) E differs from a F_σ set by a set of measure 0

countable union of closed sets.

Proof: (ii) $\Rightarrow$ (i) and (iii) $\Rightarrow$ (i) are immediate.

(i) $\Rightarrow$ (ii) E measurable $\Rightarrow$ $\forall n \geq 1$, $\exists$ open $U_n \supseteq E$ s.t. $m(U_n \setminus E) \leq \frac{1}{n}$

If $U = \bigcap_{n=1}^{\infty} U_n$, then U is G_δ and $m(U \setminus E) = 0$.

(i) $\Rightarrow$ (iii) Same thing with closed sets contained in E.

$\square$

So, we know how big $\mathcal{M}$ has to be. But how big *can* it be?

Q: Is every set measurable?

A: No. Consider $[0,1] \subseteq \mathbb{R}$.

Define an equivalence relation $\sim$ on $[0,1]$ by $x \sim y \Leftrightarrow x-y \in \mathbb{Q}$

Then, $\sim$ partitions $[0,1]$ into equivalence classes $[0,1] = \bigcup_{\alpha \in I} \mathcal{E}_\alpha$ $\leftarrow$ $\sim$-equivalence classes

Take $x_\alpha \in \mathcal{E}_\alpha$ and $\underline{\text{form the set}}$ $N := \{x_\alpha : \alpha \in I\}$

forming this set relies on the axiom of choice!

Enumerate $\mathbb{Q} \cap [-1,1]$ as $\{r_n : n \in \mathbb{N}\}$ and consider for each $n \geq 1$, $N_n = N + r_n$.

Then, $\{N_n\}_{n \in \mathbb{N}}$ are disjoint, since if two N_n's differed by a rational, we selected two representatives from the same equivalence class, which we did not.

So, $\forall x \in [0,1]$, $x \in \mathcal{E}_\alpha$ for some α, which means that x differs from an x_α by some rational $r_n \in \mathbb{Q} \cap [-1,1]$, which means $x \in N_n$ for some n.

$$\Rightarrow [0,1] \subseteq \bigcup_{n=1}^{\infty} N_n \subseteq [-1,2] \quad (*)$$

If N measurable, then N_n measurable $\forall n$ and $m(N) = m(N_n)$ because they are translates. By $(*)$,

$$m([0,1]) \leq m\left(\bigcup_{n=1}^{\infty} N_n\right) \leq m([-1,2]) \Rightarrow 1 \leq \sum_{n=1}^{\infty} m(N) \leq 3$$

So, $m(N)$ cannot be 0 and $m(N)$ cannot be >0. $\divideontimes$, so N isn't measurable. $\square$

§1.4 - Measurable Functions

Aim: Find a notion of functions we can integrate.

The simplest kind of function is an indicator function.

Def: The characteristic function of a set $E \subseteq \mathbb{R}^n$ is $\mathbf{1}_E(x) = \chi_E(x) = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}$

The next simplest kind of function is a finite sum of indicators.

Def: A simple function is a function of the form

$$\sum_{i=1}^{N} a_i \mathbf{1}_{E_i}, \text{ where } a_i \in \mathbb{R} \text{ and } E_i \text{ are measurable with finite measure.}$$

Remark: Recall Riemann integration is defined with step fns, which are $\sum_{i=1}^{N} a_i R_i$ for rectangles R_i. Simple fns. are more general.

We consider functions $f: \mathbb{R}^n \to [-\infty, \infty]$. We say f is finite-valued if $-\infty < f(x) < \infty$ $\forall x \in \mathbb{R}^n$.

Remark: Most fns. we consider are finite valued almost everywhere.

We want to form Lebesgue integration by multiplying level set values by the measure of the preimage of the value, and so we want preimages to be measurable.

Defn: If E is measurable and $f: E \to [-\infty, \infty]$, then f is a measurable function if $\forall a \in \mathbb{R}$, $f^{-1}([-\infty, a))$ is measurable.

Notation: $f^{-1}([-\infty, a)) = \{x \in E : f(x) < a\} = \{f < a\}$

Remark: we could use $[-\infty, a]$ or other stuff. This is equivalent because we can reframe things with unions, intersections, and complements, which measure behaves well under.

In a sense, we are requiring that preimages of Borel sets are measurable. More generally, we might look at preimages of elements of a certain σ-algebra being elements of a σ-algebra.

Proposition 3: (Properties of Measurable Functions)

1. If f is finite valued, then f is measurable $\Leftrightarrow f^{-1}(U)$ measurable $\forall U$ open.
   (to remove finite-valued, also assume $f^{-1}(\{-\infty\})$, $f^{-1}(\{+\infty\})$ measurable)
2. If $f: \mathbb{R}^n \to \mathbb{R}$ is continuous, it is measurable.
3. If f is measurable, finite-valued, and Φ is continuous, then $\Phi \circ f$ is measurable.
   (not true if f is not finite valued)
4. If $\{f_k\}_{k=1}^\infty$ are measurable, then so are
   $$\sup_k f_k, \quad \inf_k f_k, \quad \limsup_k f_k, \quad \liminf_k f_k$$
5. If $\{f_k\}_{k=1}^\infty$ measurable and converge $f_k \to f$ pointwise, then f is measurable.
6. If f, g measurable, then so are
   (i) f^k for $k \in \mathbb{N}$
   (ii) $f+g$ and fg if f, g finite-valued

Lecture 2/13-

Def: Two functions $f,g: E \to \mathbb{R}$ agree almost everywhere (a.e.) if $\{x \in E : f(x) \neq g(x)\}$ has measure 0.

Prop: If f, g agree a.e. and f is measurable, then so is g.

Proof: $\{f > a\}$ and $\{g > a\}$ differ by a set of measure 0.

So, $\{f > a\}$ measurable $\Rightarrow \{g > a\}$ measurable. $\square$

Remark: Because of the above, all properties from Prop. 3 hold if you replace equality with equality a.e.

Theorem 4.1- Suppose $f: \mathbb{R}^n \to [0, \infty]$ is non-negative measurable. Then, $\exists$ an increasing sequence of simple functions $(\varphi_k)_{k=1}^{\infty}$ converging to f pointwise everywhere.

(measurable fns. are good candidates for integration)

Proof: Fix $N \in \mathbb{N}$. Set $Q_n = [-N, N]^n$ to be the cube of side length $2N$. Set

$$F_N(x) := \mathbb{1}_{Q_N}(x) \cdot \min\{f(x), N\}.$$

So, we truncated to domain Q_N and range $[0, N]$.
By Prop. 3, F_N is measurable. This converges to f pointwise as $N \to \infty$.
Now, subdivide the range further. Fix $M \in \mathbb{N}$, let

preimage of a slice $\longrightarrow$ $$E_{N,j} := \left\{x \in E : \frac{j}{M} < F_N(x) \leq \frac{j+1}{M}\right\} \quad \text{for } j = 0, \dots, NM-1$$

Each $E_{N,j}$ is measurable and, since each $E_{N,j} \subseteq Q_N$, each $E_{N,j}$ has finite measure. Set

$$\varphi_{N,M} = \sum_{j=0}^{NM-1} \frac{j}{M} \mathbb{1}_{E_{N,j}}$$

This is a simple function and $\varphi_{N,M} \leq F_N$. Also, $|F_N - \varphi_{N,M}| \leq \frac{1}{M}$ on Q_N.
Now set $N = M = 2^k$ for $k \in \mathbb{N}$. Take $\varphi_k := \varphi_{2^k, 2^k} \Rightarrow |F_{2^k} - \varphi_k| \leq \frac{1}{2^k}$.

So, since $F_{2^k} \to f$ pointwise and $\varphi_k \to F_{2^k}$ in norm, then $\varphi_k \to f$ pointwise and is an increasing set of simple functions. $\square$

$|\varphi_k - f| \leq |\varphi_k - F_{2^k}| + |F_{2^k} - f| \leq 2\varepsilon$

We can now use this to remove the non-negativity assumption!

Theorem 4.2: Suppose $f: \mathbb{R}^n \to [-\infty, \infty]$ is measurable. Then, $\exists$ a sequence $(\varphi_k)_{k=1}^\infty$ of simple functions with $\varphi_k \to f$ pointwise and $|\varphi_k(x)| \le |\varphi_{k+1}(x)|$ $\forall k, x$.

Proof: Split f into positive and negative parts $f = f^+ - f^-$, where $f^+(x) := \max\{f(x), 0\}$ and $f^-(x) := -\min\{f(x), 0\}$. This is the oldest trick in measure theory! Since f^+ and f^- are both measurable and ≥ 0, Theorem 4.1 gives $(\varphi_k^+)_{k=1}^\infty$, $(\varphi_k^-)_{k=1}^\infty$ of increasing simple functions with $\varphi_k^+ \to f^+$ pointwise.
$\varphi_k^- \to f^-$

Set $\varphi_k = \varphi_k^+ - \varphi_k^-$ to get $\varphi_k \to f$, and

$$|\varphi_k| = |\varphi_k^+ - \varphi_k^-| = |\varphi_k^+| + |\varphi_k^-| \le |\varphi_{k+1}^+| + |\varphi_{k+1}^-| = |\varphi_{k+1}|$$

this holds since only one of them is ever nonzero.

$\square$

Theorem 4.3: Suppose $f: \mathbb{R}^n \to [-\infty, \infty]$ is measurable. Then, $\exists$ a sequence $(\psi_k)_{k=1}^\infty$ of step functions with $\psi_k \to f$ pointwise a.e.

Sketch proof: Theorem 4.2 yields $(\varphi_k)_k$ s.t. $\varphi_k \to f$ everywhere. Recall from ages ago that if we had a set of finite measure, we can find finitely many rectangles s.t. the symmetric difference has small measure (Prop. 3.4(iv)). So, we can find a step fn. ψ_k s.t. $\psi_k = \varphi_k$ on some measurable set F_k^c, where $m(F_k) < \frac{1}{2^k}$ (small symmetric difference).
Now, we can Borel-Cantelli on

$$F = \limsup_{k \to \infty} F_k := \bigcap_{\ell=1}^{\infty} \bigcup_{k \ge \ell} F_k \overset{\text{B.C.}}{\Longrightarrow} m(F) = 0.$$

If $x \notin F$, then $\exists \ell$ s.t. $\forall k \ge \ell$, $x \in F_k^c \Rightarrow \psi_k(x) = \varphi_k(x) \xrightarrow[k \to \infty]{} f(x)$

$\square$

We now know that measurable functions are limits of sequences of simple functions everywhere and of step functions a.e. We can now integrate!

We saw before that measurable sets aren't "too funky" as they differ from G_δ or F_σ by sets of measure 0.

Two questions for measurable functions:

1. How different are pointwise and uniform convergence
2. How different are measurable functions from continuous functions?

Answer to ①:

Theorem: (Egorov's Theorem)

Suppose $(f_k)_{k=1}^{\infty}$ are measurable, defined on a measurable set E of finite measure. Suppose $f_k \to f$ pointwise a.e. on E.

Then, $\forall \varepsilon > 0$, $\exists$ a closed $A_\varepsilon \subseteq E$ s.t. $m(E \setminus A_\varepsilon) < \varepsilon$ and $f_k \to f$ uniformly on A_ε.

Proof: WOLOG, assume $f_k \to f$ everywhere. $\forall n, k \in \mathbb{N}$, set

$$E_{n,k} := \{x \in E : |f_\ell(x) - f(x)| < \tfrac{1}{n} \ \forall \ell \geq k\}$$

For fixed n, $(E_{n,k})_{k=1}^{\infty}$ are increasing. By pointwise convergence, they increase to E.

$\Longrightarrow E \setminus E_{n,k}$ decreases to $\emptyset$. $m(E) < \infty \Rightarrow \lim_{k \to \infty} m(E \setminus E_{n,k}) = 0$.

For each n, we can then choose k_n s.t. $m(E \setminus E_{n,k_n}) < \frac{1}{2^n}$.

For $\varepsilon > 0$, choose N s.t. $\sum_{n=N}^{\infty} \frac{1}{2^n} < \varepsilon$.

Set $\tilde{A}_\varepsilon := \bigcap_{n \geq N} E_{n,k_n} \Rightarrow m(E \setminus \tilde{A}_\varepsilon) \leq \sum_{n=N}^{\infty} m(E \setminus E_{n,k_n}) < \sum_{n=N}^{\infty} \frac{1}{2^n} \leq \varepsilon$.

We claim $f_k \to f$ uniformly on $\tilde{A}_\varepsilon$. To see this, fix $\delta > 0$. Choose an $n_* \geq N$ s.t. $\frac{1}{n_*} < \delta$. Then, $x \in \hat{A}_\varepsilon \Rightarrow x \in E_{n_*, k_{n_*}} \Rightarrow |f_\ell(x) - f(x)| \leq \frac{1}{n_*} < \delta \ \forall \ell \geq k_{n_*}$.

Since n_*, k_{n_*} are independent of x, $f_k \to f$ uniformly on $\tilde{A}_\varepsilon$.
Now find closed $A_\varepsilon \subseteq \tilde{A}_\varepsilon$ with $m(\tilde{A}_\varepsilon \setminus A_\varepsilon) < \varepsilon$.
Then, $f_k \to f$ uniformly on A_ε and $m(E \setminus A_\varepsilon) < 2\varepsilon$.

$\square$

Lecture 2/15

Answer to ②:

Theorem (Lusin's Theorem)

Suppose $f : E \to \mathbb{R}$ is finite-valued and measurable, where E is measurable with $m(E) < \infty$. Then, $\forall \varepsilon > 0$, $\exists$ a closed set $F_\varepsilon \subseteq E$ with

$$m(E \setminus F_\varepsilon) \leq \varepsilon \quad \text{and} \quad f|_{F_\varepsilon} \text{ is continuous.}$$

Remark: "$f|_{F_\varepsilon} : F_\varepsilon \to \mathbb{R}$ is continuous" is weaker than saying "f is continuous on F_ε"

For example, $f = \mathbb{1}_{[0,1] \cap (\mathbb{R} \setminus \mathbb{Q})}$ vs. $f|_{[0,1] \cap (\mathbb{R} \setminus \mathbb{Q})} \equiv 1$

not continuous — continuous

Proof: Theorem 4.3 gives that $\exists$ step functions $(S_n)_{n=1}^{\infty}$ with $S_n \to f$ pointwise a.e. Note that step functions are indicator functions of rectangles, and so are discontinuous at their boundaries.
Then, for each n we can find $E_n \subseteq E$ s.t. $S_n|_{E \setminus E_n}$ is continuous and $m(E \setminus E_n) < \frac{1}{2^n}$. (just remove neighborhood around rectangle boundary)

Fix $\varepsilon > 0$. Egorov's Theorem yields $A_\varepsilon \subseteq E$ with $m(E \setminus A_\varepsilon) \leq \varepsilon$ and $S_n \to f$ uniformly on A_ε.

Choose N s.t. $\sum_{n=N}^{\infty} \frac{1}{2^n} < \varepsilon$, and set $\tilde{F}_\varepsilon := A_\varepsilon \setminus \bigcup_{n=N}^{\infty} E_n$

We then have $m(E \setminus \tilde{F}_\varepsilon) \leq 2\varepsilon$ (one from A_ε, one from $\sum m(E \setminus E_n)$) and $S_n \to f$ uniformly on $\tilde{F}_\varepsilon$ and S_n is continuous on $\tilde{F}_\varepsilon$ $\forall n \geq N$.
Since continuity is inherited by uniform limits, $f|_{\tilde{F}_\varepsilon}$ is continuous.

Take a closed set $F_\varepsilon \subseteq \tilde{F}_\varepsilon$ with $m(\tilde{F}_\varepsilon \setminus F_\varepsilon) \leq \varepsilon$.
Then, $m(E \setminus F_\varepsilon) \leq 3\varepsilon$. $\square$

End of Chapter 1

§2: Integration Theory

§2.1: Lebesgue Integral

We will build up the integral on progressively more general functions:

(i) start w/ simple functions
(ii) bounded measurable functions on sets of finite measure
(iii) non-negative measurable functions
(iv) measurable functions

(i) - Simple Functions

Note: Simple functions don't have unique representations (you can split sets). We need to ensure that integrals are well-defined. We will use the canonical form of simple functions.

Def: A simple function $S=\sum_{i=1}^{N} a_i \mathbb{1}_{E_i}$ is in canonical form if $a_i \neq a_j$ $\forall i \neq j$, and $\{E_i\}_i$ are pairwise disjoint.

Such a form always exists: every simple S takes on finitely many distinct values, say $\tilde{a}_1, \dots, \tilde{a}_N$. Letting $\tilde{E}_i := \{x \in E : S(x) = \tilde{a}_i\}$, then

$$S = \sum_{i=1}^{\tilde{N}} \tilde{a}_i \mathbb{1}_{\tilde{E}_i}$$

Def: (Integral of Simple Functions)

If $S=\sum_{i=1}^{N} a_i \mathbb{1}_{E_i}$ is a simple function in canonical form, we define its Lebesgue integral by

$$\int_{\mathbb{R}^n} S(x)\,dx := \sum_{i=1}^{N} a_i\, m(E_i)$$

Also, for $E \subseteq \mathbb{R}^n$ measurable, we define

$$\int_E S(x)\,dx := \int_{\mathbb{R}^n} \overbrace{S(x)\cdot \mathbb{1}_E}^{\text{still a simple function}}\,dx$$

Notation: To stress Lebesgue integration, we write $\int_{\mathbb{R}^n} S(x)\,dm(x)$. We can use as shorthand $\int S$ or $\int_{\mathbb{R}^n} S$

Prop. 1.1

Lebesgue integration of simple functions obeys:

(i) if $S=\sum_{i=1}^{N} a_i \mathbb{1}_{E_i}$ is any representation of a simple function, then
$$\int_{\mathbb{R}^n} S\,dx = \sum_{i=1}^{N} a_i\, m(E_i)$$

(ii) if s_1, s_2 simple and $a,b \in \mathbb{R}$ then:
$$\int (as_1 + bs_2) = a\int s_1 + b\int s_2$$

(iii) if E, F disjoint & measurable and s simple, then
$$\int_{E\cup F} S = \int_E S + \int_F S$$

(iv) if s_1, s_2 simple and $s_1 \leq s_2$ a.e, then
$$\int s_1 \leq \int s_2$$

(v) if s simple then so is $|s|$ and
$$\left|\int s\right| \leq \int |s|$$

(vi) if s_1, s_2 simple and agree a.e., then
$$\int s_1 = \int s_2$$

Proof: Assume (i) first, and prove the rest.

(ii): follows from (i) by writing down any representation

(iii): follows from (ii), as $\mathbb{1}_{E\cup F} = \mathbb{1}_E + \mathbb{1}_F$ for disjoint E, F

(iv): if $s \geq 0$ a.e. is simple, then $s = \sum_{i=1}^{N} a_i \mathbb{1}_{E_i}$ where $m(E_i) \neq 0 \Rightarrow a_i \geq 0$

$\Rightarrow \int s = \sum_{i: a_i \geq 0} a_i m(E_i) \geq 0$. Letting $s = s_2 - s_1$, linearity yields the result.

(v): $s = \sum_{i=1}^{N} a_i \mathbb{1}_{E_i}$ in canonical form $\Rightarrow |s| = \sum_{i=1}^{N} |a_i| \mathbb{1}_{E_i}$

$$\Rightarrow \left|\int s\right| = \left|\sum_{i=1}^{\tilde{N}} a_i m(E_i)\right| \underset{\text{Triangle ineq.}}{\leq} \sum_{i=1}^{N} |a_i| m(E_i) = \int |s|$$

(vi): proof is same as (iv)

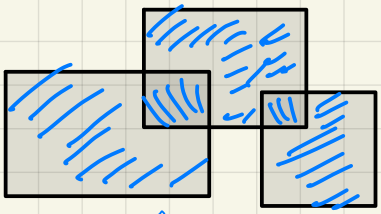
(i): Case 1: assume that $\{E_i\}_i$ are pairwise disjoint, but the a_i's could agree.
Write $\tilde{a}_1, \dots, \tilde{a}_{\tilde{N}}$ for the distinct a_i's. Set $\tilde{E}_i := \bigcup_{j: a_j = \tilde{a}_i} E_j$ for $i = 1, \dots, \tilde{N}$.
Clearly $\{\tilde{E}_i\}_i$ are pairwise disjoint and

$$s = \sum_{i=1}^{N} a_i \mathbb{1}_{E_i} = \underbrace{\sum_{i=1}^{\tilde{N}} \tilde{a}_i \mathbb{1}_{\tilde{E}_i}}_{\text{canonical form}} \Rightarrow \int s = \sum_{i=1}^{\tilde{N}} \tilde{a}_i m(\tilde{E}_i) = \sum_{i=1}^{\tilde{N}} \tilde{a}_i \left(\sum_{j: a_j = \tilde{a}_i} m(E_j)\right)$$

$$= \sum_{j=1}^{N} a_j m(E_j) \quad \text{as desired.}$$

Case 2: Now, suppose we are in the general case $s = \sum_{i=1}^{N} a_i \mathbb{1}_{E_i}$.
You can find $\{\tilde{E}_j\}_{j=1}^{\tilde{N}}$ s.t.

- $\bigcup_{i=1}^{N} E_j = \bigcup_{i=1}^{\tilde{N}} \tilde{E}_j$,
- $\{\tilde{E}_j\}_{j=1}^{\tilde{N}}$ is pairwise disjoint
- $E_i = \bigcup_{j: \tilde{E}_j \subseteq E_i} \tilde{E}_j$



$\underbrace{}_{}$ $2^N - 1$ possible $\tilde{E}_j$'s, since you don't want all complements

In fact, the $\tilde{E}_j$ are of the form $E_1' \cap \dots \cap E_N'$, where $E_i' \in \{E_i, E_i^c\}$

Now, if $\tilde{a}_j := \sum_{j: \tilde{E}_j \subseteq E_i} a_i$, then

$$s = \sum_{k=1}^{N} a_k \mathbb{1}_{E_k} = \sum_{k=1}^{\tilde{N}} a_k \mathbb{1}_{\underset{\text{disjoint union}}{\bigcup_{j: \tilde{E}_j \subseteq E_k} \tilde{E}_j}} = \sum_{k=1}^{N} a_k \sum_{j: \tilde{E}_j \subseteq E_k} \mathbb{1}_{\tilde{E}_j} = \sum_{j=1}^{\tilde{N}} \tilde{a}_j \mathbb{1}_{\tilde{E}_j}$$

$$\text{Case 1} \Rightarrow \int s = \sum_{j=1}^{\tilde{N}} \tilde{a}_j m(\tilde{E}_j) = \dots = \sum_{k=1}^{N} a_k m(E_k)$$

$\square$

We are now done with fucking with simple functions, and can treat it as a black box in the future.

Lecture 2/20 -

Now that we defined the integral for simple functions, we can proceed by integrating <u>bounded</u> measurable functions <u>supported</u> on sets of <u>finite</u> measure.

Defn: The support $f: A \to \mathbb{R}$ is
$$\mathrm{supp}(f) = \mathrm{spt}(f) = \{x \in A : f(x) \neq 0\}$$
We say f is supported on E if $f(x) = 0 \ \forall x \in E^c$

Note: f measurable $\Rightarrow$ $\mathrm{supp}(f)$ measurable

Theorem 4.2 gave that if f measurable, $|f| \leq M$, and supported on E, then $\exists (\varphi_n)_{n=1}^{\infty}$ simple s.t. $|\varphi_n| \leq M$, $\mathrm{supp}(\varphi_n) \subseteq E$, and $\varphi_n \to f$ pointwise.
(boundedness & support come from the fact that (φ_n) is increasing)

Lemma 1.2:

Suppose $f: \mathbb{R}^n \to \mathbb{R}$ measurable with $|f| \leq M$ $(M > 0)$ and $\mathrm{supp}(f) \subseteq E$, where $m(E) < \infty$. Then, for <u>any</u> sequence of $(\varphi_n)_{n=1}^{\infty}$ simple with $|\varphi_n| \leq M$, $\mathrm{supp}(\varphi_n) \subseteq E$, and $\varphi_n \to f$ pointwise a.e., we have

- $\lim_{n \to \infty} \int \varphi_n$ exists
- the limit is independent of the sequence

Proof: Fix $\varepsilon > 0$. $m(E) < \infty$ allows us to use Egorov $\Rightarrow \exists A_\varepsilon \subseteq E$ measurable with $m(E \setminus A_\varepsilon) < \varepsilon$ and $\varphi_n \to f$ uniformly on A_ε.

$$\Rightarrow \exists N \text{ s.t. } \forall m, n \geq N, \qquad |\varphi_n(x) - \varphi_m(x)| < \varepsilon \quad \forall x \in A_\varepsilon \qquad \text{(uniformly Cauchy)}$$

So, the properties of integrating simple functions give

$$\left|\int \varphi_n - \int \varphi_m\right| = \left|\int \varphi_n - \varphi_m\right| \leq \int |\varphi_n - \varphi_m| \overset{\mathrm{supp}(\cdot) \subseteq E}{=} \int_E |\varphi_n - \varphi_m| = \int_{A_\varepsilon} \underbrace{|\varphi_n - \varphi_m|}_{< \varepsilon \text{ on } A_\varepsilon} + \int_{E \setminus A_\varepsilon} \underbrace{|\varphi_n - \varphi_m|}_{\leq |\varphi_n| + |\varphi_m| \leq 2M}$$
$$\leq \int_{A_\varepsilon} \varepsilon + \int_{E \setminus A_\varepsilon} 2M = \varepsilon \underbrace{m(A_\varepsilon)}_{\leq m(E)} + 2M \underbrace{m(E \setminus A_\varepsilon)}_{< \varepsilon} \leq \varepsilon \underbrace{(m(E) + 2M)}_{< \infty, \text{ doesn't depend on } \varepsilon}$$

The sequence $\left(\int \varphi_n\right)_{n=1}^{\infty} \subseteq \mathbb{R}$ is Cauchy $\overset{\mathbb{R} \text{ complete}}{\Longrightarrow}$ it converges!

(this idea: split into two sets, one of small measure and one on which you understand f)

For uniqueness, suppose $(\varphi_n)_n$, $(\psi_n)_n$ are two such sequences. Then,
$$\varphi_n - \psi_n \to f - f = 0 \quad \text{pointwise.}$$

Also, $\varphi_n - \psi_n$ is supported on E, and $|\varphi_n - \psi_n| \leq 2M$, so it's bounded.
By the previous reasoning, $\lim_{n\to\infty} \int \varphi_n - \psi_n$ exists.

We want to show that $f=0 \Rightarrow$ this limit is 0, and then we're done.
By the same argument, if $(\underbrace{\zeta_n}_{\varphi_n-\psi_n})$ has $\zeta_n \to 0$ pointwise a.e., Egorov gives

$\exists A_\varepsilon \subseteq E$ s.t. $m(E\setminus A_\varepsilon) < \varepsilon$ and $\zeta_n \to 0$ uniformly on A_ε

$\Rightarrow \exists N$ s.t. $\forall n \geq N$, $|\zeta_n(x)| \leq \varepsilon$ on A_ε

$$\Rightarrow \left|\int \zeta_n\right| \leq \int_E |\zeta_n| \leq \int_{A_\varepsilon} \varepsilon + \int_{E\setminus A_\varepsilon} |\zeta_n| \leq \varepsilon m(E) + \varepsilon \tilde{M} \to 0.$$

$\Rightarrow \int \zeta_n \to 0.$

$\square$

Defn: The Lebesgue integral of any bounded measurable f supported on a set of finite measure is

$$\int_{\mathbb{R}^n} f(x)\,dx := \lim_{n\to\infty} \int_{\mathbb{R}^n} \varphi_n(x)\,dx$$

for any sequence of $(\varphi_n)_{n=1}^\infty$ simple with $|\varphi_n| \leq M$, $\mathrm{supp}(\varphi_n) \subseteq C$, and $\varphi_n \to f$ pointwise a.e.

Also, for any measurable $E \subseteq \mathbb{R}^n$, define $\int_E f = \int_{\mathbb{R}^n} f \cdot \mathbb{1}_E$

By the def. of limits, our properties of integrals of simple functions apply here!

Prop. 1.3: Properties (ii)→(vi) of Prop. 1.1 are true for integration of bounded measurable functions supported on sets of finite measure.

Proof: Duh

Theorem 1.4 (Bounded Convergence Theorem)

Suppose $(f_n)_{n=1}^\infty$ are measurable, all bounded by the same $M>0$, and all supported on E with $m(E)<\infty$. Then, if $f_n \to f$ pointwise a.e., then

- f is measurable
- $|f| \leq M$ a.e.
- f is supported on E a.e.
- $\int |f_n - f| \to 0$ as $n\to\infty$ ← the important part!

Remark: $\int |f_n - f| \to 0 \Rightarrow \int f_n \to \int f \iff \lim_{n\to\infty} \int f_n = \int f = \int \lim_{n\to\infty} f_n$

Under these assumptions, we can exchange limits and integrals!

Proof: We already know f is measurable. The fact $|f| \le M$ and $\operatorname{supp}(f) \subseteq E$ a.e. follow from $f_n \to f$ pointwise a.e.

Fix $\varepsilon > 0$. By Egorov, $\exists A_\varepsilon \subseteq E$ measurable with $m(E \setminus A_\varepsilon) < \varepsilon$ and $f_n \to f$ uniformly on A_ε. So,

$$\int |f_n - f| = \int_E |f_n - f| = \int_{A_\varepsilon} |f_n - f| + \int_{E \setminus A_\varepsilon} |f_n - f| \le \varepsilon m(E) + 2M\varepsilon \to 0$$

$\square$

Note: We already knew uniform convergence $\Rightarrow \lim \int = \int \lim$. Egorov shows uniform convergence except on a set of small measure. Decreasing this along the sequence yields the result.

Note: If $f \ge 0$ a.e. and f measurable, and $\int f = 0$, then $f = 0$ a.e.
To see this, set $\tilde{f} = \mathbb{1}_{B_R(0)} \cdot \min\{f, 1\}$.

Note that $\operatorname{supp}(\tilde{f}) \subseteq B_R(0)$ and $|\tilde{f}| \le 1$ and $\tilde{f} \le f$ on $B_R(0)$.
So, $\int \tilde{f} \le \int f = 0 \Rightarrow \int \tilde{f} = 0$. But $\forall k \ge 1$, $\underbrace{\frac{1}{k} \mathbb{1}_{\{\tilde{f} > \frac{1}{k}\}} \le \tilde{f}}_{\text{useful trick!}}$

$$\Rightarrow \int \frac{1}{k} \mathbb{1}_{\{\tilde{f} > \frac{1}{k}\}} \le \int \tilde{f} = 0$$

$$\Rightarrow \frac{1}{k} m(\{\tilde{f} > \tfrac{1}{k}\}) = 0 \ \forall k \Rightarrow m(\{\tilde{f} > 0\}) = m\left(\bigcup_{k=1}^{\infty} \{\tilde{f} > \tfrac{1}{k}\}\right) = 0$$

$$\Rightarrow \tilde{f} = 0 \text{ a.e.} \Rightarrow f = 0 \text{ a.e. in } B_R(0) \underset{R \to \infty}{\Rightarrow} f = 0 \text{ a.e.}$$

$\square$

Riemann Integration & Lebesgue Integration

We can now prove: if $f : [a,b] \to \mathbb{R}$ is Riemann integrable, then

- f is measurable
- the Riemann and Lebesgue integrals agree

Riemann integrable $\Rightarrow f$ bounded. $\operatorname{Domain}(f) = [a,b] \Rightarrow m(\operatorname{supp}(f)) < \infty$.
Also, one can find sequences of step functions $(\varphi_k)_{k=1}^{\infty}, (\psi_k)_{k=1}^{\infty}$ s.t.

- $(\varphi_k)_k, (\psi_k)_k$ are uniformly bounded
- $(\varphi_k)_k$ increasing, $(\psi_k)_k$ decreasing, and $\varphi_k \le f \le \psi_k \ \forall k$
- $\lim_{k \to \infty} \int_{[a,b]}^{\mathcal{R}} \varphi_k = \int_{[a,b]}^{\mathcal{R}} f = \lim_{k \to \infty} \int_{[a,b]}^{\mathcal{R}} \psi_k$ (Riemann integrals)

Since Riemann and Lebesgue integration agree on step functions, $\int_{[a,b]}^{\mathcal{R}} \varphi_k = \int_{[a,b]}^{\mathcal{L}} \varphi_k$.

Since ψ_k decreasing and bounded below by f, they converge pointwise. Same with φ_k.
Let φ, ψ be the pointwise limits. Clearly, $\varphi \le f \le \psi$.

Bounded Convergence $\Rightarrow$ $\lim_{k\to\infty}\int^{\mathcal{L}}_{[a,b]}\varphi_k = \int^{\mathcal{L}}_{[a,b]}\varphi$, $\lim_{k\to\infty}\int^{\mathcal{L}}_{[a,b]}\psi_k = \int^{\mathcal{L}}_{[a,b]}\psi$ $\Rightarrow \int^{\mathcal{L}}_{[a,b]}\psi = \int^{\mathcal{R}}_{[a,b]} f = \int^{\mathcal{L}}_{[a,b]}\varphi$

$\Rightarrow \int^{\mathcal{L}}_{[a,b]} \overbrace{\psi-\varphi}^{\geq 0} = 0 \Rightarrow \psi-\varphi=0$ a.e. $\Rightarrow \psi=\varphi$ a.e. $\Rightarrow f = \underbrace{\psi=\varphi}_{\text{measurable}} \Rightarrow f$ measurable

$\Rightarrow \varphi_k \to f$ pointwise a.e. $\overset{\text{Bounded convergence}}{\Longrightarrow} \int^{\mathcal{L}}_{[a,b]}\varphi_k \to \int^{\mathcal{L}}_{[a,b]} f \Longrightarrow \int^{\mathcal{R}}_{[a,b]} f = \int^{\mathcal{L}}_{[a,b]} f$

$= \lim_{k\to\infty}\int^{\mathcal{R}}_{[a,b]}\varphi_k = \int^{\mathcal{R}}_{[a,b]} f$

(Note: the two limits agreeing is because $\mathbb{R}$ is a Hausdorff space)

Now, we can define integration on <u>any</u> non-negative measurable function.

<u>Def:</u> For $f:\mathbb{R}^n\to[0,\infty]$ measurable, we define its (extended) Lebesgue integral by

$$\int_{\mathbb{R}^n} f(x)\,dx := \sup_g \int_{\mathbb{R}_n} g(x)\,dx$$

where the sup is taken over all g measurable with $0\leq g\leq f$ where g is bounded and supported on a set of finite measure.

We say f is (Lebesgue) integrable if this sup is finite.

As always, if $E\subseteq\mathbb{R}^n$ measurable, $\int_E f(x)\,dx = \int_{\mathbb{R}^n} f(x)\cdot \mathbb{1}_E\,dx$

<u>Prop. 1.6</u>: The integral of nonnegative measurable functions obeys:

(i) if $f,g\geq 0$, $a,b\in[0,\infty)$, then $\int af+bg = a\int f + b\int g$

(ii) if E,F disjoint and $f\geq 0$, then $\int_{E\cup F} f = \int_E f + \int_F f$

(iii) if $0\leq f\leq g$, then $\int f\leq \int g$

(iv) if $0\leq f\leq g$ with g integrable, then f integrable

(v) if $f\geq 0$ integrable, then $f(x)<\infty$ a.e.

(vi) if $f\geq 0$ and $\int f=0$, then $f=0$ a.e.

Lecture 2/22 -

Proof:

(i) Recall that $\int f := \sup_g \int g$ ($0 \le g \le f$, g bounded with $\mathrm{supp}(g)$ finite measure). We see $\int af = a\int f$ simply by rescaling each $g \le af$. So, suppose WLOG that $a=b=1$.

($\ge$) Take $f_1, g_1 \ge 0$ bounded, supports finite measurable, with $f_1 \le f$, $g_1 \le g \Rightarrow 0 \le f_1 + g_1 \le f+g$ and $f_1 + g_1$ will still be bounded and on support of finite measure ($\mathrm{supp}(f_1+g_1) \subseteq \mathrm{supp}(f_1) + \mathrm{supp}(g_1)$) $\Rightarrow \int (f_1+g_1) \le \int (f+g) \Rightarrow \int f_1 + \int g_1 \le \int (f+g)$

Taking the sup over all such f_1, g_1, $\int f + \int g \le \int (f+g)$

($\le$) Take any $0 \le \mathfrak{z} \le f+g$ bounded and supported on finite measure.
Consider $f_1 := \min\{f, \mathfrak{z}\} \Rightarrow 0 \le f_1 \le f$ and $f_1 \le \mathfrak{z} \Rightarrow f_1$ bounded and $\mathrm{supp}(f_1) \subseteq \mathrm{supp}(\mathfrak{z})$

and $g_1 := \mathfrak{z} - f_1 \Rightarrow \underbrace{0 \le g_1 \le g}_{\text{since } f_1 \text{ is either } \mathfrak{z} \text{ or } f}$ and g_1 bounded + support finite measure

$$\Rightarrow \int \mathfrak{z} = \int (f_1 + g_1) = \int f_1 + \int g_1 \underset{\text{by def. of integral}}{\le} \int f + \int g$$

Taking the sup over $\mathfrak{z} \le f+g$, $\int (f+g) \le \int f + \int g$.

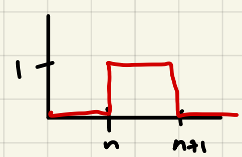
(ii) - (v) come naturally from the definition.

$\square$

Q: We saw earlier the Bounded Convergence Theorem. Can we remove either the boundedness or finite support conditions?

It would be really nice if $f_n \ge 0$ and $f_n \to f$ pointwise a.e. $\Rightarrow \int f_n \to \int f$.
This is not true: here are two examples:

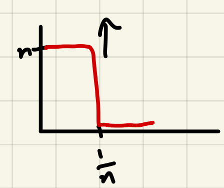
① $f_n = \mathbb{1}_{[n, n+1]} \Rightarrow f_n \to 0$ pointwise, but $\int f_n = 1 \not\to 0 = \int f$



$\{f_n\}$ has a support of $[0, \infty)$ (pushing mass to ∞)

② $f_n = n\,\mathbb{1}_{[0, \frac{1}{n}]} \Rightarrow f_n \to 0$ pointwise a.e., but $\int f_n = 1 \not\to 0 = \int f$

$\{f_n\}$ is not uniformly bounded (local blowup of measure)



(These are pretty much the two things that can go wrong with measurable functions!)

However, we do have the following:

Super important!

☆ Lemma (Fatou's Lemma)

Suppose $(f_n)_{n=1}^{\infty}$, $f_n \geq 0$, and $f_n \to f$ pointwise a.e.

Then, $$\int f \leq \liminf_{n\to\infty} \int f_n \Rightarrow \int \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int f_n$$

Proof: Take any $0 \leq g \leq f$ bounded and with support of finite measure. Consider

$$g_n := \min\{g, f_n\} \Rightarrow 0 \leq g_n \leq f_n,\ g_n \text{ bounded, and } supp(g_n) \subseteq supp(g)$$

Note that $g_n \to g$ pointwise a.e.. Bounded Convergence gives $\int g_n \to \int g$

Then,
$$\int g_n \leq \int f_n \Rightarrow \liminf_{n\to\infty} \int g_n \leq \liminf_{n\to\infty} \int f_n \underset{\int g_n \to \int g}{\Rightarrow} \int g \leq \liminf_{n\to\infty} \int f_n$$

Since $g \leq f$ arbitrary, taking the sup over g yields $\int f \leq \liminf_{n\to\infty} \int f_n$.

$\square$

☆ Corollary 1.8

If f, $(f_n)_n$ are measurable and non-negative, $f_n \to f$ pointwise a.e., and $f_n \leq f$, then $$\int f_n \to \int f$$

Proof: $f_n \leq f \Rightarrow \int f_n \leq \int f \Rightarrow \limsup_{n\to\infty} \int f_n \leq \int f \underset{\text{Fatou}}{\leq} \liminf_{n\to\infty} \int f_n$

However, $\liminf \leq \limsup \Rightarrow \lim_{n\to\infty} \int f_n$ exists and equals $\int f$.

$\square$

☆ Corollary 1.9: (Monotone Convergence Theorem)

If $f_n \geq 0$ measurable and $f_n \nearrow f$ pointwise a.e., then
$$\int f_n \to \int f.$$

Proof: Apply Corollary 1.8 as $f_n \leq f$ $\forall n$.

$\square$

☆ Corollary: (Exchanging Infinite Sums w/ Integrals)

If $a_k \geq 0$ are measurable, then
$$\int \sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} \int a_k$$

Proof: Take $f_n := \sum_{k=1}^{n} a_n \Rightarrow f_n \nearrow f = \sum_{k=1}^{\infty} a_k$. Apply Monotone Convergence. $\square$

Note: In the above, if $\sum_{k=1}^{\infty} \int a_k < \infty$, then the above gives $\int \sum_{k=1}^{\infty} a_k < \infty$
$\Rightarrow \sum_{k=1}^{\infty} a_k < \infty$ converges a.e.

We can now do the general case of integration!

Def: If $f: \mathbb{R}^n \to [-\infty, \infty]$ measurable, we say f is (Lebesgue) integrable if $|f| \geq 0$ is integrable, as defined earlier. (i.e. $\int |f| < \infty$)

When f is integrable, we define its (Lebesgue) integral by

$$\int f := \int f^+ - \int f^- \quad , \text{ where } \quad \begin{array}{l} f^+ := \max\{f, 0\} \geq 0 \\ f^- := \max\{-f, 0\} \geq 0 \end{array}$$

Remarks:

① as redefining f on a set of measure 0 doesn't change $\int f$, we allow f to be undefined on a set of measure 0.

② Since f integrable $\Rightarrow f$ finite valued a.e., we can add integrable functions as the only ambiguity in the sum is still on a set of measure 0.

③ For these reasons, we essentially are talking about equivalence classes of functions under $f \sim g \Leftrightarrow f = g$ a.e.

Prop. 1.1: The integral of integrable functions is linear, additive, monotone, and satisfies the Triangle inequality.

Proof: Follows from def. and non-negative case. $\square$

Aside: if $f: \mathbb{R}^n \to \mathbb{C}$, $f = u + iv$ with $u, v: \mathbb{R}^n \to \mathbb{R}$.
We say f is integrable if $|f| = \sqrt{u^2 + v^2}$ is integrable.
Since $|u|, |v| \leq |f|$ and $|f| \leq |u| + |v|$, then

$$f \text{ integrable} \Leftrightarrow u, v \text{ integrable}$$

Naturally, we define

$$\int f = \int u + i \int v$$

With a general integral definition, we can go ahead with:

Theorem 1.13: (Dominated Convergence)

Suppose $(f_n)_n$ measurable, and $f_n \to f$ pointwise a.e.

Then, if $\exists$ a single integrable g with $|f_n| \leq g$ a.e. $\forall n$, we have

$$\int |f_n - f| \to 0 \quad \Rightarrow \quad \int f_n \to \int f$$

Lecture 2/27

Proof of dominated convergence: Set $E_k := \{x : |x|, g(x) \leq k\}$ for $k > 0$.

Then, $g \cdot \mathbb{1}_{E_k} \nearrow g$ are increasing non-negative fns. So, if we fix $\varepsilon > 0$, monotone convergence gives $\exists k > 0$ s.t. $\int g - \int g \cdot \mathbb{1}_{E_k} < \varepsilon \Rightarrow \int_{E_k^c} g < \varepsilon$. Fix this $k > 0$.

Consider $\{f_n \cdot \mathbb{1}_{E_k}\}_n$. Each one is bounded and on a set of finite measure.

Bounded convergence gives $\int_{E_k} |f_n - f| \to 0 \Rightarrow \exists N$ s.t. $\forall n \geq N$, $\int_{E_k} |f_n - f| < \varepsilon$

But for these n,

$$\int |f_n - f| = \int_{E_k} \overbrace{|f_n - f|}^{\leq \varepsilon} + \int_{E_k^c} \overbrace{|f_n - f|}^{\leq |f_n| + |f| \leq 2g} \leq \varepsilon + 2\int_{E_k^c} g \leq 3\varepsilon.$$

$\square$

As a corollary, we prove (not in the book!) the following:

Theorem: (Differentiating under the integral sign)

Suppose $U \subseteq \mathbb{R}$ (our parameter, like $t \rightsquigarrow \frac{\partial}{\partial t}$) and $f : U \times \mathbb{R}^n \to \mathbb{R}$ is s.t.

(i) $x \to f(t,x)$ is integrable $\forall t$

(ii) $t \mapsto f(t,x)$ is differentiable $\forall x$, with continuous derivative

(iii) $\exists g$ integrable with $\left|\frac{\partial f}{\partial t}(t,x)\right| \leq g$ for a.e. x, t.

Then, $x \mapsto \frac{\partial f}{\partial t}(t,x)$ is integrable $\forall t$, and the map $F(t) = \int_{\mathbb{R}^n} f(t,x)\,dx$ is differentiable with $F'(t) = \int \frac{\partial f}{\partial t}(t,x)\,dx$.

In other words, $\frac{\partial}{\partial t} \int_{\mathbb{R}^n} f(t,x)\,dx = \int_{\mathbb{R}^n} \frac{\partial f}{\partial t}(t,x)\,dx$.

Proof: $x \to \frac{\partial f}{\partial t}(t,x)$ integrable is clear from domination.

Take $(h_k)_k \subseteq \mathbb{R}$ with $h_k \to 0$, and look at

$$g_k(x) := \frac{f(t+h_k, x) - f(t,x)}{h_k} - \frac{\partial f}{\partial t}(t,x) > 0 \quad \text{for } t \text{ fixed.}$$

The MVT gives $g_k(x) = \frac{\partial f}{\partial t}(\tilde{t}, x) - \frac{\partial f}{\partial t}(t,x)$ for some $\tilde{t} \in [t, t+h_k]$

$\Rightarrow |g_k| \le |\underbrace{2g}_{\text{integrable}}|$ $\forall k$. Continuity of $\frac{\partial f}{\partial t} \Rightarrow g_k \to 0$ pointwise.

We can apply dominated convergence to get $\int g_k \to \int 0 = 0$

But
$$\int g_k = \frac{F(t+h_k) - F(t)}{h_k} - \int_{\mathbb{R}^n} \frac{\partial f}{\partial t}(t,x)\,dx$$

So,
$$\frac{F(t+h_k) - F(t)}{h_k} \longrightarrow \int_{\mathbb{R}^n} \frac{\partial f}{\partial t}(t,x)\,dx$$

$\Rightarrow F$ differentiable with $F'(t) = \int_{\mathbb{R}^n} \frac{\partial f}{\partial t}(t,x)\,dx$

$\square$

Prop 1.12:

Suppose $f: \mathbb{R}^n \to \mathbb{R}$ is integrable. Then, $\forall \varepsilon > 0$,

(i) $\exists R > 0$ s.t. $\int_{B_R^c(0)} |f| < \varepsilon$ ← no mass at ∞

(ii) $\exists \delta > 0$ s.t. $m(E) < \delta \Rightarrow \int_E |f| < \varepsilon$ ← no mass accumulation on small sets

these are the two things that can go wrong!

Proof: (i) We've seen $|f| \cdot \mathbb{1}_{B_R(0)} \nearrow |f|$ as $R \to \infty$ allows monotone convergence

$$\Rightarrow \int |f| \cdot \mathbb{1}_{B_R(0)} \to \int |f| \Rightarrow \int |f| \cdot \mathbb{1}_{B_R^c(0)} \to 0$$

(ii) $|f| \cdot \mathbb{1}_{\{|f| \le N\}} \nearrow |f|$ as $N \to \infty$. Monotone convergence implies that

$$\exists N \text{ s.t. } \int |f| - \int |f| \cdot \mathbb{1}_{\{|f| \le N\}} < \varepsilon \Rightarrow \int |f| \cdot \mathbb{1}_{\{|f| > N\}} < \varepsilon$$

So, if $m(E) < \delta$,
$$\int_E |f| = \int_E \underbrace{|f| \cdot \mathbb{1}_{\{|f| > N\}}}_{< \varepsilon} + \int_E \underbrace{|f| \cdot \mathbb{1}_{\{|f| \le N\}}}_{\le N} \le \varepsilon + N\delta$$

N was fixed, so selecting $\delta = \frac{\varepsilon}{N}$ gives us the result.

$\square$

§2.2: The L^1 Space

We've seen that Lebesgue integrable functions form a vector space. With the right $\|\cdot\|$, it forms a <u>complete</u> normed vector space.

<u>Def:</u> The vector space of (equivalence classes of) Lebesgue integrable functions $f: \mathbb{R}^n \to \mathbb{R}$ forms a normed vector space when endowed with the norm

$$\|f\|_{L^1} := \int_{\mathbb{R}^n} |f| \, dx$$

This space is called $L^1(\mathbb{R}^n)$.

<u>Some properties:</u>

(i) $\|af\|_{L^1} = |a| \, \|f\|_{L^1} \quad \forall f \in L^1, a \in \mathbb{C}$

(ii) $\|f+g\|_{L^1} \leq \|f\|_{L^1} + \|g\|_{L^1} \quad \forall f, g \in L^1$

(iii) $\|f\|_{L^1} = 0 \iff f = 0$ a.e. (f is in 0's equivalence class)

(iv) $d(f,g) := \|f-g\|_{L^1}$ is a metric on L^1

This can be generalized.

<u>Def:</u> Let $p \in [1, \infty)$. The vector space of (equivalence classes of) measurable functions $f: \mathbb{R}^n \to \mathbb{R}$ with $\int |f|^p < \infty$ forms a normed vector space when endowed with the norm

$$\|f\|_{L^p} := \left(\int |f|^p \right)^{\frac{1}{p}}$$

This space is called $L^p(\mathbb{R}^n)$.

<u>Remark:</u> An L^p space is dual to an L^q space with $\frac{1}{p} + \frac{1}{q} = 1$. Since Hilbert spaces are dual to themselves, only L^2 can be Hilbert.

Theorem: (Riesz-Fischer)

$L^1(\mathbb{R}^n)$ is **complete** (i.e. every Cauchy sequence converges)

Proof: Suppose $(f_n)_n \subseteq L^1(\mathbb{R}^n)$ is Cauchy. Consider any subsequence $(f_{n_k})_k$ with

$$\|f_{n_{k+1}} - f_{n_k}\|_{L^1} \le 2^{-k}$$ (can be done by spamming Cauchy criterion with $\varepsilon_k = 2^{-k}$)

Look at
$$f(x) := f_{n_1}(x) + \sum_{k=1}^{\infty}\left(f_{n_{k+1}}(x) - f_{n_k}(x)\right)$$ and
$$g := |f_{n_1}| + \sum_{k=1}^{\infty}|f_{n_{k+1}} - f_{n_k}|$$

Clearly $|f| \le g$, and

$$\int g = \int |f_{n_1}| + \int \sum_{k=1}^{\infty}|f_{n_{k+1}} - f_{n_k}| \overset{\text{monotone converge}}{=} \int |f_{n_1}| + \sum_{k=1}^{\infty}\int |f_{n_{k+1}} - f_{n_k}| \le \int |f_{n_1}| + \sum_{k=1}^{\infty} 2^{-k} < \infty$$

So, g is integrable $\Rightarrow g \in L^1 \Rightarrow |f| < \infty$ a.e. $\Rightarrow f$ in L^1.

However, since f is a telescoping sum. So, $\lim_{k\to\infty} f_{n_k}(x)$ exists a.e. with $f(x)$ as the pointwise limit a.e.

To upgrade to $f_{n_k} \to f$ in L^1, note that $\underbrace{|f - f_{n_k}|}_{\text{subtracting off a partial sum}} \le g \quad \forall k$

Dominated convergence yields $\int |f - f_{n_k}| \to 0 \Rightarrow f_{n_k} \to f$ in L^1.

Since subsequence convergence $\Rightarrow$ sequence convergence for Cauchy, we are done.

$\square$

Lecture 3/1

Corollary 2.3:

If $(f_n)_n \subseteq L^1(\mathbb{R}^n)$ and $f_n \to f$ in L^1, then $\exists$ a subsequence $(f_{n_j})_j$ with $f_{n_j} \to f$ pointwise a.e.

Proof: $f_n \to f$ in $L^1 \Rightarrow (f_n)_n$ Cauchy in L^1, so we can apply the first half of the previous proof.

$\square$

Def: A subset $A \subseteq L^1(\mathbb{R}^n)$ is *dense* if $\forall f \in L^1(\mathbb{R}^n)$ and $\varepsilon > 0$, $\exists g \in A$ s.t. $\|f - g\|_{L^1} < \varepsilon$.

Theorem 2.4: The following subsets of $L^1(\mathbb{R}^n)$ are dense:

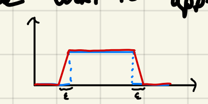
(i) the simple functions
(ii) the step functions
(iii) $C_c(\mathbb{R}^n)$ - the continuous functions which have compact support

technically, can do same thing with cont., infinitely diff'able fns

Proof: WOLOG, by approximating real/imaginary parts separately, suppose functions are real-valued. Also, WOLOG, by splitting $f = f^+ + f^-$ and approximating separately, suppose functions are ≥ 0.

(i) Theorem 4.1 from Chap. 1 $\Rightarrow$ simple functions dense.

(ii) All we must show is that step functions are dense in simple functions, and the result then follows from (i). So, all we must show is that step functions approximate $\mathbb{1}_E$ for any measurable E with $m(E) < \infty$. Theorem 4.3 (iv) from Chap. 1 $\Rightarrow \exists$ closed rectangles $(R_i)_{i=1}^N$ with $m(E \,\Delta \bigcup_i R_i) < \varepsilon$. The step fn $\sum_{i=1}^N \mathbb{1}_{R_i}$ then works in the sense that $\|\mathbb{1}_E - \sum_{i=1}^N \mathbb{1}_{R_i}\|_{L^1} < \varepsilon$.

(iii) We WTS $C_c(\mathbb{R}^n)$ is dense in the step functions. So, we want to approximate $\mathbb{1}_R$ for some closed rectangle R of finite measure. For $n=1$, simply  by linear interpolation.

For general n, $R = [a_1, b_1] \times \dots \times [a_n, b_n]$, so take a product of the 1D fns above. $\square$

Remark: To prove things about L^1, prove about a dense subset and pass the property through a limit.

Note: From translational and scaling invariance of Lebesgue measure, we can show through simple functions that

- $\int_{\mathbb{R}^n} f(x-h)\,dx = \int_{\mathbb{R}^n} f(x)\,dx \quad \forall h \in \mathbb{R}^n$
- $\int_{\mathbb{R}^n} f(ax)\,dx = \frac{1}{a^n}\int_{\mathbb{R}^n} f(x)\,dx \quad \forall a > 0$
- $\int_{\mathbb{R}^n} f(-x)\,dx = \int_{\mathbb{R}^n} f(x)\,dx$

If we write $f_h(x) := f(x-h)$ for $h \in \mathbb{R}^n$, clearly $f_h \to f$ pointwise as $h \to 0$ depends on continuity of f, which isn't true $\forall f \in L^1(\mathbb{R}^n)$. However, $f_h \to f$ in L^1.

Prop. 2.5:

If $f \in L^1(\mathbb{R}^n)$, then $f_h \to f$ in $L^1(\mathbb{R}^n)$ as $h \to 0$.

Proof. Let $\varepsilon > 0$. Theorem 2.4 (iii) $\Rightarrow \exists g \in C_c(\mathbb{R}^n)$ s.t. $\|f-g\|_{L^1} < \varepsilon$. We have

$$f_h - f = (f_h - g_h) + (g_h - g) + (g - f) \Rightarrow \|f_h - f\|_{L^1} \leq \underbrace{\|f_h - g_h\|_{L^1}}_{= \|f-g\|_{L^1} < \varepsilon} + \|g_h - g\|_{L^1} + \underbrace{\|g - f\|_{L^1}}_{< \varepsilon}$$

Note that $\|g_h - g\|_{L^1} = \int_{\mathbb{R}^n} |g(x-h) - g(x)|\, dx \xrightarrow[\text{convergence}]{g, g_h \text{ compact support, continuous} \Rightarrow \text{bounded}} \|g_h - g\|_{L^1} < \varepsilon$ for some h by bdd. convergence

$\Rightarrow \|f_h - f\|_{L^1} < 3\varepsilon$, completing the proof.

$\square$

§2.3: Fubini's Theorem

Q: ① when can we swap the order of integration?
② when can you compute a higher dim. integral via separate lower-dim. integrals?

Def. Let $E \subseteq \mathbb{R}^n \times \mathbb{R}^m$ have coordinates $(x,y) \in \mathbb{R}^n \times \mathbb{R}^m$. We define slices by
$$E^y = \{x \in \mathbb{R}^n \mid (x,y) \in E\} \quad \text{and} \quad E_x := \{y \in \mathbb{R}^m \mid (x,y) \in E\}$$
If $f: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, its slices are
$$f^y(x) := f(x,y) \quad \text{and} \quad f_x(y) := f(x,y)$$

Note: We know slices of Borel sets are Borel. However, it isn't true that E measurable $\Rightarrow$ slices of E are (consider $\mathcal{N} \times \{0\}$ has measure 0 in $\mathbb{R}^2$, but $\mathcal{N}$ sucks in $\mathbb{R}$).

Note: It is not true that "f measurable $\Rightarrow f^y$ measurable". However, almost every slice is measurable.

Theorem 3.1: (Fubini)

Suppose f is integrable on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Then, for a.e. $y \in \mathbb{R}^{d_2}$,

(i) the slice f^y is integrable on $\mathbb{R}^{d_1}$

(ii) the map $y \mapsto \int_{\mathbb{R}^{d_1}} f^y(x)\, dx$ is integrable on $\mathbb{R}^{d_2}$

Moreover,
$$\int_{\mathbb{R}^{d_1 \times d_2}} f = \int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f(x,y)\, dx \right) dy$$

Swapping x,y gives
$$\int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f(x,y)\, dx \right) dy = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x,y)\, dy \right) dx$$

Proof: As usual, WOLOG f is real valued. Let $\mathcal{F}$ be the set of all integrable functions satisfying the conclusions of Fubini's theorem. We WTS $L^1(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}) \subseteq \mathcal{F}$. We perform a monotone class argument.

Step 1: If $f, g \in \mathcal{F}$, then $f+g \in \mathcal{F}$.

Proof: If A, B denote the sets of measure 0 away from which f, g obey the conclusions, then f, g obey the conclusions away from $A \cup B$, and $m(A \cup B) = 0$

integrable away from $A \cup B$

So, $(f+g)^y = f^y + g^y \Rightarrow (f+g)^y$ is integrable for $y \notin A \cup B$. Similarly,

$$\int (f+g)^y = \int f^y + \int g^y < \infty \Rightarrow f+g \in \mathcal{F}.$$

Step 2: If $(f_k)_k \subseteq \mathcal{F}$ and $f_k \nearrow f$ pointwise a.e. to some $f \in L^1(\mathbb{R}^d)$, then $f \in \mathcal{F}$.

Proof: WOLOG, $f_k \geq 0$ $\forall k$ by considering $f_k - f_1$. By monotone convergence

$$\int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f_k^y \, dx \right) dy = \int_{\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}} f_k \longrightarrow \int_{\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}} f$$

If f_k^y integrable $\forall y \notin A_k$ with $m(A_k) = 0$, then f_k^y integrable $\forall k$, $\forall y \notin \underbrace{\bigcup_{k=1}^{\infty} A_k}_{\text{measure } 0} := A$

Monotone convergence gives $\int_{\mathbb{R}^{d_1}} f_k^y \to \int_{\mathbb{R}^{d_1}} f^y \quad \forall y \notin A$.

So,

$$\int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f_k^y \, dx \right) dy \to \int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f^y \, dx \right) dy$$

But

$$\int_{\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}} f < \infty \Rightarrow \int_{\mathbb{R}^{d_2}} \left(\int_{\mathbb{R}^{d_1}} f^y \right) < \infty \Rightarrow \int_{\mathbb{R}^{d_1}} f^y < \infty \text{ for a.e. } y.$$

So, f^y integrable for a.e. y.

Lecture 3/6-

At this point in the proof, we have that $\mathcal{F}$ is closed under finite linear combinations and pointwise monotone limits

Step 3: If E is a G_δ set with finite measure, then $\mathbb{1}_E \in \mathcal{F}$.

Proof: We build up from simple sets

- if $E = Q_1 \times Q_2$, then every slice of E is measurable (it's either $\emptyset$ or a cube)
  open cubes in $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$
  and moreover $\int_{\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}} \mathbb{1}_E = |Q_1||Q_2| = \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} \mathbb{1}_E \, dy \, dx$
- if $E \subseteq$ boundary of some cube, then a.e. slice is empty (or at least measurable) (boundary is measure 0) and all the integrals are $0 \Rightarrow$ Fubini holds $\Rightarrow \mathbb{1}_E \in \mathcal{F}$.

- if E = finite union of almost disjoint cubes, we can write E as a disjoint union of interiors and boundaries. By the above two bullets and Step 1, $\mathbb{1}_E \in \mathcal{F}$.

- if E is open with finite measure, then Theorem 1.4 (§1.1) $\Rightarrow E = \bigcup_{j=1}^{\infty} Q_j$ countable union of almost disjoint cubes. Since $\mathbb{1}_{\bigcup_{j=1}^{N} Q_j} \in \mathcal{F}$ and $\mathbb{1}_{\bigcup_{j=1}^{N} E_j} \nearrow \mathbb{1}_E$, Step 2 gives that $\mathbb{1}_E \in \mathcal{F}$.

- if E is any G_δ set of finite measure, $E = \bigcap_{j=1}^{\infty} U_j$ countable intersection of open sets. Take any open set $U \supseteq E$ with finite measure. Then,

$$\mathbb{1}_E = \lim_{N \to \infty} \mathbb{1}_{U \cap \bigcap_{j=1}^{N} U_j} \quad \text{decreasing} \quad \underset{\text{Step 2}}{\Longrightarrow} \quad \mathbb{1}_E \in \mathcal{F}.$$

Step 4: If E has measure 0, then $\mathbb{1}_E \in \mathcal{F}$. We can find a G_δ-set $G \supseteq E$ with $m(G) = 0$. Step 3 says that $\mathbb{1}_G$ obeys Fubini. So,

$$\int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} \mathbb{1}_G \right) = \int_{\mathbb{R}^{d_1 \times d_2}} \mathbb{1}_G = m(G) = 0$$

Since $\int_{\mathbb{R}^{d_2}} \mathbb{1}_G(x,y)\,dy = \int_{\mathbb{R}^{d_1}} \mathbb{1}_{G_x}$ is positive and integrates to 0, $m(G_x) = 0$ for a.e. x.

So, $E_x \subseteq G_x \Rightarrow m(E_x) = 0$ for a.e. x. Then,

$$\int_{\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}} \mathbb{1}_E = m(E) = 0 \quad \text{and} \quad \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} \mathbb{1}_E(x,y)\,dy \right) dx = \int_{\mathbb{R}^{d_1}} \underbrace{m(E_x)}_{=0 \text{ a.e.}}\,dx = 0$$

This shows $\mathbb{1}_E \in \mathcal{F}$.

Step 5: Steps 1,3,4 yield that if E is measurable with finite measure, $\mathbb{1}_E \in \mathcal{F}$.

Step 6: Steps 1 + 5 gives that all simple functions are in $\mathcal{F}$.

Recall Theorem 4.1 (§1) stated that all pos. integrable functions are increasing limits of integrable functions $\underset{\text{Step 2}}{\Longrightarrow}$ all pos. integrable functions are in $\mathcal{F}$.

Since $\forall f \in L^1$, $f = f^+ - f^-$ for pos. integrable f^+, f^-, Step 1 gives that $f \in \mathcal{F}$.

$\square$

Remark: Fubini's Theorem (e.g. swapping integrals) is always true for nonnegative measurable functions (as long as the equality is understood that it could be $\infty = \infty$). This is Theorem 3.2. The proof is essentially define $f_k := f \cdot \mathbb{1}_{\{|(x,y)| < k,\ |f(x,y)| < k\}} \Rightarrow f_k \nearrow f$. Apply Fubini to each f_k, and use monotone convergence.

The usefulness is as follows. For general f measurable:

(1) check if $f \in L^1 \Leftrightarrow \int_{\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}} |f| < \infty$

(2) We can use Fubini on $|f|$ since its positive!

The above remark means we can apply Fubini to $\mathbb{1}_E$.

Corollary 3.3:

Suppose $E \subseteq \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ is measurable. Then,

(i) a.e. slice E^y, E_x is measurable

(ii) the map $y \mapsto m(E^y)$ is a measurable function, and
$$m(E) = \int_{\mathbb{R}^{d_2}} m(E^y)\, dy$$

Note: ① In general, not every slice is measurable!
e.g. $N \times \{0\} \subseteq \mathbb{R}^2$ has measure 0, but the slice at $y=0$ is N.

② If $E = E_1 \times E_2$ where E, E_1, E_2 measurable, then $m(E^y) = m(E_1)$ for $y \in E_2$ and so
$$m(E) = \int_{E_2} m(E_1)\, dy = m(E_1) m(E_2)$$

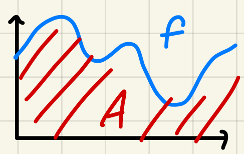
In fact, E_1, E_2 measurable $\Rightarrow E_1 \times E_2$ measurable by simple coverings by cubes (see Prop. 3.5-3.6)

This discussion also proves:

(i) if $f: \mathbb{R}^n \to \mathbb{R}$ is measurable, then $\tilde{f}: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ defined by $\tilde{f}(x,y) = f(x)$ is also measurable. This is because
$$\{\tilde{f} < a\} = \underbrace{\{f < a\}}_{\text{measurable}} \times \mathbb{R}^m \Rightarrow \{\tilde{f} < a\} \text{ measurable}$$

(ii) if $f: \mathbb{R}^n \to \mathbb{R}$ is measurable, then $\tilde{f}(x,y) = f(x-y)$ is measurable on $\mathbb{R}^{2n}$.

Corollary 3.8



Suppose $f: \mathbb{R}^n \to [0,\infty]$ and set $A = \{(x,y) \in \mathbb{R}^n \times \mathbb{R} : 0 \leq y \leq f(x)\}$
Then,

(i) f is measurable $\Leftrightarrow$ A is measurable

(ii) if it's measurable, then $\int_{\mathbb{R}^n} f = m(A)$

Proof:

(i) $(\Rightarrow)$ Let $F(x,y) := y - f(x)$. Then, f measurable $\overset{\text{above discussion}}{\Longrightarrow}$ F measurable.

$$A = \underbrace{\{y \geq 0\} \cap \{F \leq 0\}}_{\text{measurable}} \Rightarrow A \text{ measurable}$$

($\Leftarrow$) If A is measurable, then Corollary 3.3 gives that $x \mapsto m(A_x)$ is measurable, but $A_x = [0, f(x)] \Rightarrow m(A_x) = f(x)$. So, f sends $x \mapsto m(A_x)$ and f is thus measurable.

(ii) Fubini gives
$$m(A) = \int_{\mathbb{R}^n} \overbrace{m(A_x)}^{f(x)}\, dx = \int f$$

$\square$

End of Chapter 2

§3: Integration & Differentiation

There are two natural questions:

① If $f : [a,b] \to \mathbb{R}$ is integrable, then write
$$F(x) := \int_a^x f(t)\, dt$$
Is F differentiable, and if so when is $F' = f$ (a.e.)?

② If $f : [a,b] \to \mathbb{R}$, what conditions ensure that f' exists a.e., and moreover
$$\int_a^b f'(t)\, dt = f(b) - f(a) \ ?$$
Note that the Cantor-Lebesgue fn. had $f' = 0$ a.e., but $f(1) - f(0) = 1$.

From Riemann integration, we know that ① is true when f is continuous and ② is true when f is C^1.

§3.1: Differentiation of the Integral

if f continuous, take $h < \delta \Rightarrow \frac{1}{h}\int_x^{x+h} f \in f(x) \pm \varepsilon$

Look at the quotient
$$\frac{F(x+h) - F(x)}{h} = \frac{\int_a^{x+h} f(t)\, dt - \int_a^x f(t)\, dt}{h} = \frac{1}{h}\int_x^{x+h} f(t)\, dt$$

Writing $I = [x, x+h]$, we seek
$$= \frac{1}{|I|}\int_I f(t)\, dt$$

This leads us to a more general setup. In general, in $\mathbb{R}^n$, we can ask whether
$$\lim_{\substack{m(B) \to 0 \\ B \text{ ball}, \\ x \in B}} \frac{1}{m(B)} \int_B f(y)\, dy \overset{?}{=} f(x) \qquad \text{or, more generally} \qquad \lim_{\substack{m(E) \to 0 \\ E \text{ measurable}, \\ x \in E}} \frac{1}{m(E)} \int_E f(y)\, dy \overset{?}{=} f(x)$$

clearly not true. take E is some random ball union $\{x\}$. or, take $E =$, which isolates a whole segment.

So, the question is only interesting when E isolates x in its limit.

Def: Suppose $f \in L^1(\mathbb{R}^n)$. The Hardy-Littlewood maximal function of f, denoted f^*, is

$$f^*(x) = \sup_{\substack{\text{balls } B \\ x \in B}} \frac{1}{m(B)} \int_B |f(y)|\,dy$$

(the worst average you can get)

(you'd hope this is $\approx f$)

Theorem: If $f \in L^1(\mathbb{R}^n)$, then

(i) f^* is measurable (ii) $f^* < \infty$ a.e. (iii) $m(\{f^* > \alpha\}) \leq \frac{\|f\|_{L^1}}{\alpha} \cdot 3^n$

if we ignore the 3^n and replace f^* with f, this is Chebyshev. So, this is a weak L^1 estimate

Proof:

(i) $\{f^* > \alpha\}$ is open:

if $x \in \{f^* > \alpha\}$, $f^*(x) > \alpha \Rightarrow \exists$ ball B s.t. $x \in B$ and $\frac{1}{m(B)} \int_B |f(y)|\,dy > \alpha$

Suppose WOLOG that B open. Then, $\exists \delta > 0$ s.t. $B_\delta(x) \subseteq B$.
If $z \in B_\delta(x)$, then $z \in B \Rightarrow f^*(z) \geq \frac{1}{|B|} \int_B |f| > \alpha$

(ii) This follows from (iii).

$$\{f^* = \infty\} \subseteq \{f^* > \alpha\} \quad \forall \alpha > 0 \Rightarrow m(\{f^* = \infty\}) \leq \frac{1}{\alpha} \|f\|_{L^1} \cdot 3^n \quad \forall \alpha > 0$$
$$\Rightarrow m(\{f^* = \infty\}) = 0.$$

(iii) Let $x \in \{f^* > \alpha\}$. Then, as before, $\exists$ open ball B_x s.t. $x \in B_x$ and $\frac{1}{|B_x|} \int_{B_x} |f| > \alpha$

We want logic along the lines that

$$\{f^* > \alpha\} \subseteq \bigcup_{x \in \{f^* > \alpha\}} B_x \Rightarrow m(\{f^* > \alpha\}) \leq \sum_x m(B_x) \leq \frac{1}{\alpha} \sum_x \int_{B_x} |f|$$

B_x's overlap, so we need a covering lemma to deal with this

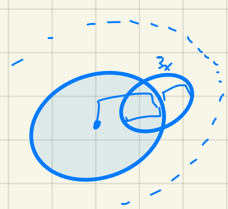
Lemma: (Vitali Covering Lemma, Elementary)

If $\mathcal{B} = \{B_1, \dots, B_N\}$ is a finite collection of open balls in $\mathbb{R}^n$, then $\exists$ disjoint subcollection $B_{i_1}, \dots, B_{i_k} \in \mathcal{B}$ with

$$m\left(\bigcup_{i=1}^N B_i\right) \leq 3^n \sum_{j=1}^k m(B_{i_j})$$

Proof of Lemma:

Intuition: consider two overlapping balls B_1, B_2 (suppose WOLOG B_1 is bigger).
Then, the ball of radius $3r(B_1)$ concentric with B_1 covers them all



For a ball B, write $\tilde{B}$ for the concentric ball with 3 times the radius. Take B_i to be the ball in $\mathcal{B}$ of largest radius. Set

$$\mathcal{B}' := \{B : B \in \mathcal{B} \text{ and } B \cap B_i \neq \emptyset\}$$

Then, $B \in \mathcal{B}' \Rightarrow B \subseteq \tilde{B}_i$. Now, throw away $\mathcal{B}'$ from $\mathcal{B}$ and consider $\mathcal{B} \setminus \mathcal{B}'$. Inductively repeat this: it terminates in finite time because each iteration removes one. Let $B_{i_1}, \dots, B_{i_K}$ be the chosen balls at each stage. Since each $B \in \mathcal{B}$ was thrown away at some point, $\exists j$ s.t. $B \subseteq \tilde{B}_{i_j} \Rightarrow \bigcup_{B \in \mathcal{B}} B \subseteq \bigcup_{j=1}^{K} \tilde{B}_{i_j}$

$$\Rightarrow m\Big(\bigcup_{B\in\mathcal{B}} B\Big) \leq \sum_{j=1}^{K} m(\tilde{B}_{i_j}) = 3^n \sum_{j=1}^{K} m(B_{i_j})$$

□

Note that for any compact $K \subseteq \{f^* > \alpha\}$, $K \subseteq \bigcup_x B_x \xrightarrow{\text{finite subcover}} K \subseteq \bigcup_{i=1}^{N} B_i$
Apply the lemma to this collection of balls. Then,

$$\Rightarrow m(K) \leq m\Big(\bigcup_{i=1}^{N} B_i\Big) \leq 3^n \sum_{j=1}^{K} m(B_{i_j}) \leq 3^n \cdot \frac{1}{\alpha} \sum_{j=1}^{K} \int_{B_{i_j}} |f| \underset{B_{i_j}\text{'s disjoint}}{=} \frac{3^n}{\alpha} \int_{\bigcup_{j=1}^{K} B_{i_j}} |f| \leq \frac{3^n}{\alpha} \int_{\mathbb{R}^n} |f|$$

As $K \subset \{f^* > \alpha\}$ is arbitrary compact set, take the sup over all such K to get (iii).

□

<u>Theorem 1.3</u>: (Lebesgue Differentiation Theorem)

If $f \in L^1(\mathbb{R}^n)$, then $\displaystyle\lim_{\substack{m(B)\to 0 \\ B \text{ ball} \\ B \ni x}} \frac{1}{m(B)} \int_B f(y)\,dy = f(x)$ for a.e. x

<u>Proof:</u> Fix $\alpha > 0$. Consider

$$E_\alpha := \left\{ x : \limsup_{\substack{m(B)\to 0 \\ B \text{ ball} \\ B \ni x}} \left| \frac{1}{m(B)} \int_B f(y)\,dy - f(x) \right| > 2\alpha \right\}$$

If we can show $m(E_\alpha) = 0$, we are done.

Fix $\varepsilon > 0$. We know (Theorem 2.4 of §2) that $\exists g$ continuous with $\|f-g\|_{L^1} < \varepsilon$. Since g continuous, we know $\displaystyle\lim_{\substack{m(B)\to 0 \\ B \text{ ball} \\ B \ni x}} \frac{1}{m(B)} \int_B g = g(x)$ for <u>all x</u>.

Now,

$$\frac{1}{m(B)} \int_B f - f(x) = \overbrace{\frac{1}{m(B)} \int_B (f-g)}^{\text{controlled by } \sup (f-g)^*} + \overbrace{\frac{1}{m(B)} \int g - g(x)}^{\text{goes to 0 since } g \text{ continuous}} + g(x) - f(x)$$

$$\Rightarrow \limsup_{\substack{m(B)\to 0 \\ B \text{ ball} \\ B \ni x}} \left| \frac{1}{m(B)} \int_B f - f(x) \right| \leq (f-g)^*(x) + 0 + |f(x) - g(x)|$$

Hence, $x \in E_\alpha \Rightarrow$ one of $(f-g)^*(x) > \alpha$ or $|f(x)-g(x)| > \alpha$

$\Rightarrow E_\alpha \subseteq F_\alpha \cup G_\alpha$ where $F_\alpha := \{(f-g)^* > \alpha\}$ and $G_\alpha := \{|f-g| > \alpha\}$

We have:

- $m(F_\alpha) \le \frac{\|f-g\|_{L^1}}{\alpha} 3^n$ by Maximal function estimate (iii),

- $m(G_\alpha) \le \frac{\|f-g\|_{L^1}}{\alpha}$ by Chebyshev

$$\Rightarrow m(E_\alpha) \le \frac{(3^n+1)}{\alpha} \varepsilon$$

Taking $\varepsilon \to 0$, we get $m(E_\alpha) = 0$.

$\square$

Lecture 3/22 -

Remarks: ① $f^*(x) \ge |f(x)|$ a.e.

② $f \in L^1(\mathbb{R}^n)$ is a global property, but Lebesgue differentiation is local. In fact, we only need to assume f is locally integrable. For example, $f(x) := x$ clearly admits the Lebesgue Differentiation Theorem, but it is not $\in L^1$.

Def: A function $f: \mathbb{R}^n \to \mathbb{R}$ is locally integrable if $\forall K \subseteq \mathbb{R}^n$ compact,
$$\int_K |f| < \infty$$
We say that $f \in L^1_{loc}(\mathbb{R}^n)$ for such f.

Consequences

① Def: If $E \subseteq \mathbb{R}^n$ is measurable, $x \in \mathbb{R}^n$ is a Lebesgue point of E if
$$\lim_{\substack{m(B) \to 0 \\ B \text{ ball} \\ B \ni x}} \frac{m(E \cap B)}{m(B)} = 1$$

Consider applying LDT with $f := \mathbb{1}_E \in L^1_{loc}(\mathbb{R}^n)$, we get

Corollary 1.5: If $E \subseteq \mathbb{R}^n$ is measurable, then

(i) a.e. $x \in E$ is a Lebesgue point of E

(ii) a.e. $x \notin E$ is not a Lebesgue point of E

Remark: this isn't that important, but it just says measurable sets are nice and don't lose mass in many places

② **Def:** If $f\in L^1_{loc}(\mathbb{R}^n)$, the Lebesgue set of f is all $x\in\mathbb{R}^n$ s.t.

(i) $|f(x)|<\infty$ (ii) $$\lim_{\substack{m(B)\to 0\\ B \text{ ball}\\ B\ni x}} \frac{1}{m(B)}\int_B |f(y)-f(x)|\,dy = 0$$

Stronger claim than result of LDT

Corollary 1.6: (Improved LDT)

If $f\in L^1_{loc}(\mathbb{R}^n)$, then a.e. $x\in\mathbb{R}^n$ is in the Lebesgue set of f.

Proof: For each $r\in\mathbb{Q}$, if we apply LDT to $|f(x)-r|\in L^1_{loc}$, we get $\exists E_r$ with measure 0 and

$$\lim_{\substack{m(B)\to 0\\ B \text{ ball}\\ B\ni x}} \frac{1}{m(B)}\int_B |f(y)-r|\,dy = |f(x)-r| \qquad \forall x\notin E_r$$

Set $E:=\bigcup_{r\in\mathbb{Q}} E_r \Rightarrow m(E)=0$. Fix $\varepsilon>0$.

If $x\notin E$ and $|f(x)|<\infty$ (this happens a.e. since $f\in L^1_{loc}$), then $\exists r\in\mathbb{Q}$ s.t. $|f(x)-r|<\varepsilon$

$$\Rightarrow \frac{1}{m(B)}\int_B |f(y)-f(x)|\,dy \le \frac{1}{m(B)}\int_B |f(y)-r|\,dy + \overbrace{\frac{1}{m(B)}\int_B |f(x)-r|\,dy}^{=|f(x)-r|<\varepsilon}$$

$$\le \frac{1}{m(B)}\int_B |f(y)-r|\,dy + \varepsilon$$

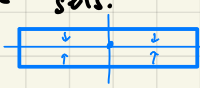
Taking limsup,

$$\limsup_{\substack{m(B)\to 0\\ B \text{ ball}\\ B\ni x}} \frac{1}{m(B)}\int_B |f(y)-f(x)|\,dy \le |f(x)-r|+\varepsilon < 2\varepsilon$$

Taking $\varepsilon\to 0$ (and noting that limsup = liminf since it's nonnegative and limsup $=0$), the result holds.

$\square$

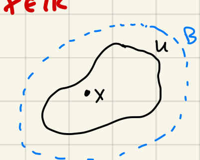
Since Lebesgue points have better averaging properties than usual points, we can use them to extend the sequences of allowed sets.

(We want to exclude skinny rectangles  and the like)

Def: A collection of measurable sets $\{U_\alpha\}_\alpha$ shrinks regularly to $x\in\mathbb{R}^n$ (or has bounded eccentricity at x) if

$\exists c>0$ s.t. $\forall U_\alpha$, $\exists$ a ball B_α with $x\in B_\alpha$ and

$U_\alpha\subseteq B_\alpha$ and $m(U_\alpha)\ge c\, m(B_\alpha)$



$m(U)\approx m(B)$ up to a multiplicative constant

Remark: {cubes} have bdd. eccentricity, but in general {rectangles} don't.

Corollary 1.7:

Suppose $f \in L^1_{loc}(\mathbb{R}^n)$. Then, if $\{U_\alpha\}_\alpha$ shrinks regularly to x and x is a Lebesgue point of f, then

$$\lim_{\substack{m(U_\alpha)\to 0 \\ U_\alpha \ni x}} \frac{1}{m(U_\alpha)} \int_{U_\alpha} f(y)\,dy = f(x)$$

Proof:

$$\left| \frac{1}{m(U_\alpha)} \int_{U_\alpha} f(y) - f(x)\,dy \right| \le \frac{1}{Cm(B)} \int_B |f(y) - f(x)|\,dy \to 0$$

as x was a Lebesgue point of f.

□

Remark: Because a.e. x is in the Lebesgue set of $f \in L^1_{loc}$, often if we want to prove something holds a.e. we assume things are Lebesgue points.

§ 3.2 - Approximations to the Identity

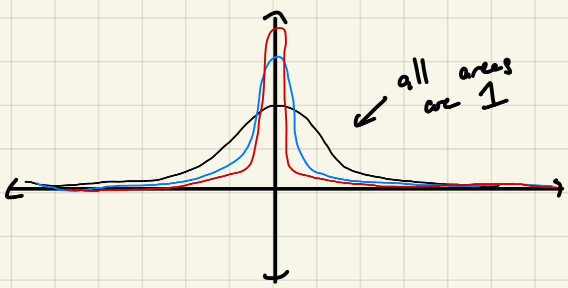
Def: An approximation to the identity K is a family functions $\{K_\delta\}_{\delta>0}$ from $\mathbb{R}^n \to \mathbb{R}$ (typically, though not always nonnegative) that obey

(integrates to 1) (i) $\int_{\mathbb{R}^n} K_\delta = 1 \quad \forall \delta$

(blows up as $\delta\to0$) (ii) $|K_\delta(x)| \le \frac{A}{\delta^n} \quad \forall \delta > 0, \forall x \in \mathbb{R}^n$, and a fixed constant A

(decays at ∞) (iii) $|K_\delta(x)| \le A\delta \cdot \frac{1}{|x|^{n+1}} \quad \forall \delta > 0, \forall x \in \mathbb{R}^n$

Picture:



The language for such objects comes from the fact (which we will prove) that $f * K_\delta$ converges in various ways to f as $\delta \to 0$.

$$(f * K_\delta)(x) = \int f(x-y)\, \overbrace{K_\delta(y)}^{\text{probability density fn.}}\, dy = \underbrace{\mathbb{E}_{a\sim K_\delta}}_{\text{as } \delta\to0,\ \mathbb{P}\{a=0\}\to 1}[f(x-a)] \to f(x)$$

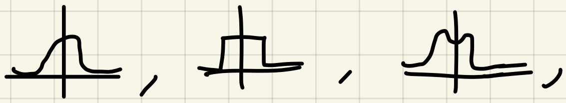
Note: $\forall r > 0$, the decay condition gives that mass concentrates at 0 via

$$\int_{|x|\ge r} |K_\delta(x)|\,dx \le A\delta \int_{|x|\ge r} \frac{1}{|x|^{n+1}}\,dx = \frac{\tilde{A}\delta}{r} \to 0 \text{ as } \delta \to 0$$

Examples:

can be achieved by normalizing

① If φ is non-negative, bounded on $\mathbb{R}^n$, with $supp(\varphi) \subseteq B$ and $\int_B \varphi = 1$, then
$$K_\delta(x) := \frac{1}{\delta^n}\varphi\left(\frac{x}{\delta}\right)$$
is an approx. to the identity.



② If you take $\varphi(x) = \frac{1}{\pi(1+x^2)}$ for $x \in \mathbb{R}$, then the $\{K_\delta\}_{\delta>0}$ generated from this are called the Poisson Kernel of the half-plane $\left(K_y(x) = \frac{y}{\pi(y^2+x^2)} \right)$

(This is how you solve the Laplacian)

③ If you take $\varphi(x) = \frac{1}{(4\pi)^{n/2}} e^{-\frac{|x|^2}{4}}$ for $x \in \mathbb{R}^n$, then the $\{K_\delta\}_{\delta>0}$ generated from this are called the heat Kernel

(This is how you solve the heat equation)

Theorem 2.1:

If $(K_\delta)_{\delta>0}$ is an approximation to the identity and $f \in L^1(\mathbb{R}^n)$, then
$$(f * K_\delta)(x) \longrightarrow f(x) \quad \text{as } \delta \to 0$$
for <u>every</u> Lebesgue point of f (in particular, $f * K_\delta \to f$ pointwise a.e.)

Note that since convolution is a sort of weighted average and Lebesgue points average nicely, this might be expected.

Proof: We have

$$(f*K_\delta)(x) - f(x) = \int_{\mathbb{R}^n} f(x-y)K_\delta(y)\,dy - \int_{\mathbb{R}^n} \underset{\text{independent of } y}{f(x)} K_\delta(y)\,dy$$
$$= \int_{\mathbb{R}^n} [f(x-y) - f(x)] K_\delta(y)\,dy$$
$$\Rightarrow |(f*K_\delta)(x) - f(x)| \leq \int_{\mathbb{R}^n} |f(x-y)-f(x)| \cdot |K_\delta(y)|\,dy$$

split $\{|y|>\delta\}$ into annuli

$$\leq \underbrace{\int_{|y|\leq\delta} |f(x-y)-f(x)| \cdot |K_\delta(y)|\,dy}_{\text{call this ①}} + \underbrace{\sum_{k=0}^{\infty} \int_{2^k\delta<|y|\leq 2^{k+1}\delta} |f(x-y)-f(x)| \cdot |K_\delta(y)|\,dy}_{\text{call this ②}}$$

Looking at ①, over this region we have $|K_\delta(y)| \le \frac{A}{\delta^n}$

$$\Rightarrow ① \le \frac{A}{\delta^n}\int_{B_\delta} |f(x-y)-f(x)|\,dy = \frac{\overset{\text{constant w.r.t. }\delta}{C}}{m(B_\delta)}\int_{B_\delta}|f(x-y)-f(x)|\,dy$$

Since x is a Lebesgue point, this $\to 0$ as $\delta\to 0$.

Lecture 3/27 -

Looking at ②,

$$\int_{2^k\delta \le |y| < 2^{k+1}\delta} |f(x-y)-f(x)|\,|K_\delta(y)|\,dy \le \int_{2^k\delta\le|y|<2^{k+1}\delta} |f(x-y)-f(x)|\cdot \frac{A\delta}{\underbrace{|y|^{n+1}}_{|y|\ge 2^k\delta}}\,dy \le \frac{A\delta}{(2^k\delta)^{n+1}}\int_{|y|<2^{k+1}\delta} |f(x-y)-f(x)|\,dy$$

annulus is smaller than whole ball, so we can make this integral be over the whole ball

$$= \frac{A\cdot 2^n}{2^k}\,\frac{1}{(2^{k+1}\delta)^n}\int_{|y|<2^{k+1}\delta}|f(x-y)-f(x)|\,dy$$

Define $\phi(r) = \frac{1}{r^n}\int_{|y|<r}|f(x-y)-f(x)|\,dy$. Our bound for all of ② is then

$$② \le A\,2^n\sum_{k=0}^{\infty} 2^{-k}\,\phi(2^{k+1}\delta)$$

As was the case in the estimate for ①, we know $\phi(r)\to 0$ as $r\to 0$.
Furthermore, $\phi(r)$ is continuous over $\mathbb{R}$. Also, $\phi(r) \le \frac{1}{r^n}\int_{|y|\le r}|f(x-y)|+|f(x)|\,dy = \frac{\|f\|_{L^1}}{r^n} + m(B_1)|f(x)|$

So, ϕ is bounded, say by some $\phi(r)\le B$.

Fix $\varepsilon>0$, choose N s.t. $\sum_{n=N}^{\infty} 2^{-n} < \varepsilon$. Now choose $\delta>0$ sufficiently small s.t. $\phi(2^{k+1}\delta) < \frac{\varepsilon}{N}$ } only depends on ε
for $k<N$.

$$\Rightarrow ② \le A2^n\left(\sum_{k=0}^{N-1}2^{-k}\phi(2^{k+1}\delta) + \sum_{k=N}^{\infty}2^{-k}\phi(2^{k+1}\delta)\right) \le A2^{n}\left(\sum_{k=0}^{N-1}1\cdot\frac{\varepsilon}{N} + \sum_{n=N}^{\infty}2^{-k}B\right)$$

$$\le A2^n(1+B)\varepsilon$$

Taking $\varepsilon\to 0$, ② $\to 0$ as $\delta\to 0$.

$\square$

One can also prove L^1 convergence.

Theorem 2.3:

Suppose $f\in L^1(\mathbb{R}^n)$ and $\{K_\delta\}_\delta$ is an approximation to the identity. Then,

(i) $\forall\delta>0$, $f*K_\delta \in L^1(\mathbb{R}^n)$
(ii) $\|(f*K_\delta)-f\|_{L^1}\to 0$ as $\delta\to 0$ (i.e. $f*K_\delta\to f$ in L^1)

Proof: on the next PSET :)

$\square$

§ 3.3 - Differentiability of Functions

We want to find broad conditions on F that ensure $F(b)-F(a) = \int_a^b F'(t)\,dt$

(Minter says this might be the hardest thing we do in the course).

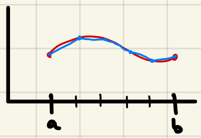
Some issues we expect:

- even if F is continuous, F' may not exist
- F' may exist a.e., but F' may not be integrable

To characterize possible F's, we want to characterize functions arising as indefinite integrals. We start by looking at functions of bounded variation (which is related to lengths of curves and other geometric things).

Def: Let $\gamma \subseteq \mathbb{R}^2$ be a curve, parameterized by $z(t) = (x(t), y(t))$ where x and y are continuous. We say γ is rectifiable if

$$L(\gamma) := \sup \sum_{j=1}^{N} |z(t_j) - z(t_{j-1})| < \infty$$



supremum over polygonal approximations to the curve

where the supremum is taken over all partitions of the domain $z: [a,b] \to \mathbb{R}^2$ given by $a = t_0 < t_1 < \dots < t_N = b$. We call $L(\gamma)$ the length of γ.

Thinking about rectifiability leads us to

Def: Suppose $F: [a,b] \to \mathbb{C} \cong \mathbb{R}^2$. Consider a partition (always finite P's) $\mathcal{P} := \{a = t_0 < t_1 < \dots < t_N = b\}$ of $[a,b]$. The variation of F w.r.t. $\mathcal{P}$ is $\sum_{j=1}^{N} |F(t_j) - F(t_{j-1})| < \infty$

We say F is of bounded variation (written $F \in BV([a,b])$) if

$$\sup_{\text{partitions } \mathcal{P}} \sum_{j=1}^{N} |F(t_j) - F(t_{j-1})| < \infty$$

Note:

- when talking about rectifiability of curves, we also assume continuity. For variation we don't.
- If $\tilde{\mathcal{P}}$ is a partition which refines $\mathcal{P}$ (contains more points), then variation w.r.t. $\tilde{\mathcal{P}} \geq$ variation w.r.t. $\mathcal{P}$

Def: We say that $F:[a,b]\to\mathbb{R}$ is

- **increasing** if $x<y \Rightarrow F(x)\le F(y)$
- **strictly increasing** if $x<y \Rightarrow F(x)<F(y)$

Examples

① If F is increasing and bounded, then it is of bounded variation as

$$\sum_j |F(t_j)-F(t_{j-1})| = \sum_j \left(F(t_j)-F(t_{j-1})\right) = F(b)-F(a)$$

② If F is Lipschitz, then F is of bounded variation. In particular, if F is differentiable <u>everywhere</u> with bounded derivative, then F is Lipschitz $\Rightarrow$ bounded variation

Def: Let $F:[a,b]\to\mathbb{R}$ be a function. Then, the **total variation** of F on $[a,x]$ with $x\in[a,b]$ is

$$T_F(a,x) := \sup_{\substack{\text{partitions of} \\ [a,x]}} \sum_{j=1}^{N} |F(t_j)-F(t_{j-1})| \ge 0$$

The **positive variation** is

$$P_F(a,x) := \sup_{\substack{\text{partitions of} \\ [a,x]}} \sum_{(+)} |F(t_j)-F(t_{j-1})| \ge 0$$

$(+) = \{j : F(t_j)\ge F(t_{j-1})\}$

The **negative variation** is

$$N_F(a,x) := \sup_{\substack{\text{partitions of} \\ [a,x]}} \sum_{(-)} |F(t_j)-F(t_{j-1})| \ge 0$$

$(-) = \{j : F(t_j)\le F(t_{j-1})\}$

Lemma 3.2:

Suppose $F\in BV([a,b])$. Then, $\forall x\in[a,b]$,

$$F(x)-F(a) = P_F(a,x) - N_F(a,x) \qquad \text{and} \qquad T_F(a,x) = P_F(a,x) + N_F(a,x)$$

the definitions give <u>two</u> partitions, but we can take the union and use monotonicity of P_F, N_F under finer partitions

Proof: Fix $\varepsilon>0$. Using the definitions of P_F, N_F as suprema, one can find a partition $a=t_0<t_1<\dots<t_n=b$ s.t.

$$\left|P_F(a,x) - \sum_{(+)} (F(t_j)-F(t_{j-1}))\right| < \varepsilon \qquad \text{and} \qquad \left|N_F - \sum_{(-)} -(F(t_j)-F(t_{j-1}))\right| < \varepsilon$$

Now, $F(x)-F(a) = \sum_j (F(t_j)-F(t_{j-1})) = \sum_{(+)} (F(t_j)-F(t_{j-1})) - \sum_{(-)} -(F(t_j)-F(t_{j-1}))$ for all partitions

$\Rightarrow |F(x)-F(a)-(P_F-N_F)| < 2\varepsilon$. Taking $\varepsilon\to 0$, the first result holds.

For the second part, note that for any partition of $[a,x]$,

$$\sum_j |F(t_j)-F(t_{j-1})| = \sum_{(+)} (F(t_j)-F(t_{j-1})) + \sum_{(-)} -(F(t_j)-F(t_{j-1})) \le P_F + N_F$$

Taking the supremum over all partitions, $T_F\le P_F+N_F$. Symmetrically,

$$\sum_{(+)} (F(t_j)-F(t_{j-1})) + \sum_{(-)} -(F(t_j)-F(t_{j-1})) = \sum_j |F(t_j)-F(t_{j-1})| \le T_F \quad \text{for all partitions}$$

Using another ε-argument as above, $P_F+N_F+2\varepsilon \le T_F \overset{\varepsilon\to 0}{\Rightarrow} P_F+N_F\le T_F$. The claim follows. $\square$

This gives:

Theorem 3.3:

Let $F:[a,b]\to\mathbb{R}$. Then,

F is of bounded variation $\iff$ $F=f_1-f_2$, where f_1 and f_2 are increasing bounded functions

Proof: ($\Leftarrow$) $f_1,f_2\in BV([a,b])$ by Example 1. The result follows.
($\Rightarrow$) Set $f_1(x):=P_F(a,x)+F(a)$, $f_2(x):=N_F(a,x)$ $\Rightarrow$ f_1,f_2 bounded since $F\in BV([a,b])$ and f_1,f_2 increasing since P_F,N_F increasing

By Lemma 3.2, $F=f_1-f_2$. □

Lecture 3/29 -

Remarks:
- can get equivalent result for $F:[a,b]\to\mathbb{C}$ or $\mathbb{R}^n$ by looking componentwise
- can also show that F continuous $\Rightarrow T_F(a,\cdot)$ is continuous.

A key result is then: (this is super duper important in solving PDEs, Sobolev spaces, etc)
(Lipschitz $\Rightarrow$ BV $\Rightarrow$ diff'able is an important foundation for geometric measure theory)

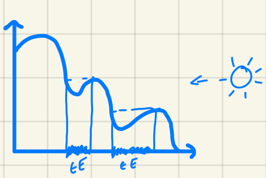
Theorem 3.4:

If $F:[a,b]\to\mathbb{R}$ is of bounded variation, then F is differentiable a.e.

Proof: First, assume that F is continuous as well.

Lemma 3.5: (Rising Sun Lemma)

Suppose $G:\mathbb{R}\to\mathbb{R}$ is continuous. Set $E:=\{x: G(x+h)>G(x) \text{ for some } h>0\}$. If $E\neq\emptyset$, then E is open in $\mathbb{R}$ (as G is continuous), and so $E=\bigcup_{k=1}^{\infty}(a_k,b_k)$ is a countable union of disjoint open intervals



Then, for any bounded (a_k,b_k) we have $G(a_k)=G(b_k)$

Proof: Look at some (a_k,b_k). We know $a_k,b_k\notin E$ since the intervals are not disjoint. So, $G(a_k)\geq G(b_k)$. Suppose BWOC that $G(a_k)>G(b_k)$. By IVT, $\exists c_k\in(a_k,b_k)$ with $G(c_k)=\dfrac{G(a_k)+G(b_k)}{2}$. Choose c to be maximal with this property (something maximal is a limit point, which must be $\leq b_k$ but can't be b_k because $G(c)>G(b_k)$).
But $c\in E\Rightarrow \exists d>c$ with $G(d)>G(c)$. But $b_k\notin E$, so $G(b_k)\geq G(\text{everything bigger than } b_k)$.
But $G(d)>G(b_k)\Rightarrow d<b_k$. But then $d<b_k$ and $G(d)>G(c)>G(b_k)$. So, IVT gives that $\exists e\in(d,b_k)$ with $G(e)=G(c)$. But $e>c$ and c was selected maximally. □

The above proof also gives

<u>Corollary 3.6:</u>

Suppose now that $G:[a,b]\to\mathbb{R}$. Then, if a_k could be a for one of the intervals, in which case all we know is that $G(a_k)\le G(b_k)$

Under the assumption that F is continuous, define $(\Delta_h F)(x) := \frac{F(x+h)-F(x)}{h}$

Consider the 4 Dini numbers:

$$D^+F(x) := \limsup_{h\downarrow 0} (\Delta_h F)(x) \qquad D_+F(x) := \liminf_{h\downarrow 0} (\Delta_h F)(x)$$

$$D^-F(x) := \limsup_{h\uparrow 0} (\Delta_h F)(x) \qquad D_-F(x) := \liminf_{h\uparrow 0} (\Delta_h F)(x)$$

Clearly, $D_+\le D^+$ and $D_-\le D^-$ as $\liminf \le \limsup$.
We <u>claim</u>:
(i) $D^+F(x) < \infty$ a.e.
(ii) $D^+F(x)\le D_-F(x)$ for a.e. x

If we have these we can conclude the proof, since (ii) with $-F(-x)$ gives $D^-F(x)\le D_+F(x)$, from which we could get $D^+ \underset{(ii)}{\le} D_- \le D^- \underset{(ii)}{\le} D_+ \le D^+ < \infty \Rightarrow D^+=D_+=D^-=D_- \Rightarrow$ differentiability a.e.

Recall that F is continuous. Suppose WOLOG that it is also <u>bounded and increasing</u> because of Theorem 3.3.
can always do this when $\in BV$

Fix $\gamma>0$ and consider $E_\gamma := \{D^+F>\gamma\}$. One can show (on a PSET eventually) that E_γ is measurable. Now apply the Rising Sun Lemma to $G(x)=F(x)-\gamma x$.
The condition $G(x+h)>G(x) \Leftrightarrow F(x+h)-F(x)>\gamma h \Leftrightarrow \frac{F(x+h)-F(x)}{h}>\gamma$.

So, $E_\gamma \subseteq \bigcup_{k=1}^{\infty}(a_k,b_k)$ disjoint open intervals where $G(a_k)\le G(b_k)$ $\forall k$.

$\Rightarrow F(b_k)-F(a_k)\ge \gamma(b_k-a_k)$. So, monotonicity of measure yields

since F increasing and we are summing over segments of the range

$$m(E_\gamma)\le \sum_k m(a_k,b_k) = \sum_k (b_k-a_k) \le \frac{1}{\gamma}\sum_k (F(b_k)-F(a_k)) \le \frac{1}{\gamma}(F(b)-F(a))$$

Taking $\gamma\to\infty$, F bounded gives $m(E_\gamma)\to 0 \Rightarrow \{D^+F=\infty\}\subseteq E_\gamma$ $\forall\gamma \Rightarrow D^+F<\infty$ a.e.
So, claim (i) is proven.

For (ii): fix $R,r>0$ s.t. $R>r$ and consider

$$E_{R,r} \equiv E = \{x : D^+F(x)>R \text{ and } D_-F(x)<r\}$$

If we can show $m(E_{R,r})=0$ $\forall R>r$, we can take a union over the rationals to cover the converse of claim (ii), and we are done. Suppose BWOC that $m(E)>0$.
First, choose an open set $U\supseteq E$ with $m(U)<m(E)\cdot\frac{R}{r}$ (it's clearly measurable).

Lecture 4/3:

U open in $\mathbb{R} \Rightarrow U = \bigcup_{n=1}^{\infty} I_n$ disjoint open intervals

Applying the Rising Sun Lemma to $G(x) := -F(-x) + r(x)$ on $-I_n$, after reflecting back we get $\bigcup_k (a_k, b_k) \subset I_n$ open disjoint with $F(b_k) - F(a_k) \leq r(b_k - a_k)$.

Now apply the Rising Sun Lemma again to $G(x) := F(x) - Rx$ on (a_k, b_k). We get $\bigcup_j (a_{k_j}, b_{k_j}) \subseteq (a_k, b_k)$ s.t. $F(b_{k_j}) - F(a_{k_j}) \geq R(b_{k_j} - a_{k_j})$

Set $U_n := \bigcup_{k,j} \underbrace{(a_{k_j}, b_{k_j})}_{\text{disjoint, open}}$. Then,

F increasing $\Rightarrow \sum_j F(b_{k_j}) - F(a_{k_j}) \leq F(b_k) - F(a_k)$

$$m(U_n) = \sum_{k,j} (b_{k_j} - a_{k_j}) \leq \frac{1}{R} \sum_{k,j} \left(F(b_{k_j}) - F(a_{k_j})\right) \leq \frac{1}{R} \sum_k \left(F(b_k) - F(a_k)\right)$$

$$\leq \frac{r}{R} \sum_k (b_k - a_k) \leq \frac{r}{R} m(I_n)$$

But $E_{R,r} \cap I_n \subseteq U_n$ by defn. of E_{Rr} and the Rising Sun Lemma. Then,

$$m(E_{Rr}) = \sum_n m(E_{Rr} \cap I_n) \leq \sum_n m(U_n) \leq \frac{r}{R} \sum_n m(I_n)$$
$$= \frac{r}{R} m(U) < \frac{r}{R} \cdot \frac{R}{r} m(E_{Rr}) = m(E_{Rr}) \quad \times$$

So, $m(E_{R,r})$ must be 0. This proves (ii), and hence the result holds for continuous, bounded, increasing functions.

$\square$

We can now prove that F' (which exists a.e.) is an L^1 function:

Corollary 3.7:

If F is increasing, continuous, then F' exists a.e., F' is nonnegative, F' is measurable, and
$$\int_a^b F'(x)\,dx \leq F(b) - F(a)$$
If F bounded, then $F' \in L^1([a,b])$.

Proof: Consider the sequence of functions $G_n(x) := \dfrac{F(x+\frac{1}{n}) - F(x)}{\frac{1}{n}} \overset{F \text{ increasing}}{\geq} 0$

We know $G_n \to F'$ pointwise a.e. $\Rightarrow F'$ nonnegative a.e. and measurable.
Since $G_n \geq 0$, Fatou's Lemma gives
$$\int_a^b F' \leq \liminf_{n \to \infty} \int_a^b G_n$$

We compute
$$\int_a^b G_n = \frac{1}{\frac{1}{n}}\int_a^b F(x+\tfrac{1}{n})\,dx - \frac{1}{\frac{1}{n}}\int_a^b F(x)\,dx$$
$$= \frac{1}{\frac{1}{n}}\int_{a+\frac{1}{n}}^{b+\frac{1}{n}} F(x)\,dx - \frac{1}{\frac{1}{n}}\int_a^b F(x)\,dx$$
$$= \underbrace{\frac{1}{\frac{1}{n}}\int_b^{b+\frac{1}{n}} F(x)\,dx}_{\to F(b)} - \underbrace{\frac{1}{\frac{1}{n}}\int_a^{a+\frac{1}{n}} F(x)\,dx}_{\to F(a)}$$
by continuity of F

$$\Rightarrow \liminf_{n\to\infty}\int_a^b G_n = \lim_{n\to\infty}\int_a^b G_n = F(b)-F(a)$$
$\square$

Remark: The Cantor Lebesgue function $F:[0,1]\to[0,1]$ was continuous, bounded, increasing, and with $F(0)=0$, $F(1)=1$. But, also $F'=0$ a.e.
$\Rightarrow \int_a^b F' \neq F(b)-F(a)$. So, Corollary 3.7 can't be equality without more assumptions.

Absolute Continuity

For $f\in L^1([a,b])$, consider $F(x)=\int_a^x f(t)\,dt$.

Since $f\in L^1$, $\forall \varepsilon>0\ \exists \delta>0$ s.t. $m(E)<\delta \Rightarrow \int_E |f|<\varepsilon$

$\Rightarrow |x-y|<\delta \Rightarrow |F(x)-F(y)| = \left|\int_x^y f(t)\,dt\right| \leq \varepsilon \Rightarrow$ <u>F uniformly continuous</u>

In fact, if $(a_1,b_1),\dots,(a_N,b_N)$ are disjoint open intervals, then
$$\sum_{j=1}^N (b_j-a_j)<\delta \Rightarrow \sum_{j=1}^N |F(b_j)-F(a_j)|<\varepsilon$$
This is a stronger continuity condition known as absolute continuity.

Def: $F:[a,b]\to\mathbb{R}$ is absolutely continuous if $\forall\varepsilon>0$, $\exists\delta>0$ s.t. whenever $(a_1,b_1),\dots,(a_N,b_N)$ are disjoint with $\sum_{j=1}^N (b_j-a_j)<\delta$, then
$$\sum_{j=1}^N \left|F(b_j)-F(a_j)\right|<\varepsilon.$$

Remarks: ① integrals of L^1 functions are absolutely continuous

② absolute continuity $\Rightarrow$ uniform continuity

③ absolute continuity on $[a,b] \Rightarrow \in BV([a,b])$
In fact, total/pos/neg variation is also absolutely continuous

The main result is:

Theorem 3.8:

If F is absolutely continuous on $[a,b]$ and $F'=0$ a.e., then F is constant.

abs. cont. rules out Cantor-like behavior

Proof: Let $E=\{x: F'(x) \text{ exists and } F'(x)=0\}$; by assumption, $m(E)=b-a$.
Fix $\varepsilon>0$. If $x\in E$, we know $\lim_{h\to 0}\left|\frac{F(x+h)-F(x)}{h}\right|=0$

For all $\eta>0$, $\exists$ an interval $(a_x,b_x)=I$ containing x with
$$(b_x-a_x)<\eta \quad \text{and} \quad |F(b_x)-F(a_x)|\le\varepsilon(b_x-a_x)$$

We would like to sum these up and get variation $\le \varepsilon(b-a)$, but there could be overlaps! However, the intervals can be as small as we like (given by η), so we can use the finite version of the Vitali Covering Lemma:

Lemma 3.9: (Vitali Covering Lemma)

Suppose E is a set of finite measure. Suppose $\mathcal{B}$ is a Vitali cover (i.e. $\forall x\in E$ and $\forall\varepsilon>0$, $\exists B\in\mathcal{B}$ ball w/ $x\in B$ and $m(B)<\varepsilon$).

Then, $\forall\delta>0$, $\exists$ finitely many balls $B_1,\dots,B_N$ which are disjoint and $\sum_{i=1}^{N} m(B_i)\ge m(E)-\delta$ (almost cover E)

Proof of Lemma: Take any $\delta<m(E)<\infty$. Find a compact set $E'\subseteq E$ with $m(E')\ge\delta$. Compactness implies $E'\subseteq\bigcup_{\text{finite}} B_i$. Applying our old covering lemma, (the 3^n radius one) we find disjoint $B_1,\dots,B_{N_1}$ s.t. $3^n\sum_{i=1}^{N_1} m(B_i)\ge m(E')\ge\delta$.
If $\sum_{i=1}^{N_1} m(B_i)\ge m(E)-\delta$, we are done. Otherwise, $\sum_{i=1}^{N_1} m(B_i)<m(E)-\delta$

In this case, consider $E_2:=E\setminus\bigcup_{i=1}^{N_1}\overline{B_i}$: we know $m(E_2)\ge\delta$. The balls in $\mathcal{B}$ which are disjoint from $\bigcup_{i=1}^{N_1}\overline{B_i}$ still form a Vitali cover. So, we may repeat this argument to E_2, and so forth.
If we repeat this inductively, at each stage we throw away measure $\ge\frac{\delta}{3^n}$.
After k steps, throw away $k\delta/3^n$; as soon as $\frac{k\delta}{3^n}\ge m(E)-\delta$, we are done. $\square$

The intervals $\{(a_x,b_x)\}_{x\in E}$ forms a Vitali cover of E. Fix $\delta>0$ and apply the lemma: we get finitely many disjoint intervals $I_i:=(a_i,b_i)$ for $i=1,\dots,N$ s.t.

- $\sum_{i=1}^{N}(b_i-a_i)\ge(b-a)-\delta$
- $|F(b_i)-F(a_i)|\le\varepsilon(b_i-a_i)$

$$\Rightarrow \sum_{i=1}^{N}|F(b_i)-F(a_i)|\le\varepsilon\sum_{i=1}^{N}(b_i-a_i)\overset{\text{disjointness}}{\le}\varepsilon(b-a)$$

But now, $[a,b] \setminus \bigcup_{i=1}^{\tilde{N}} (a_i, b_i)$ is a finite union of disjoint intervals $[\alpha_j, \beta_j]$ whose total length is $< \delta$. Choosing δ appropriately, absolute continuity gives

$$\sum_{j=1}^{\tilde{m}} |F(\beta_j) - F(\alpha_j)| \le \varepsilon$$

But now $|F(b) - F(a)| \le \sum_{i=1}^{\tilde{N}} |F(b_i) - F(a_i)| + \sum_{j=1}^{\tilde{m}} |F(\beta_j) - F(\alpha_j)| \le \varepsilon(b-a) + \varepsilon \overset{\varepsilon \downarrow 0}{\Longrightarrow} F(a) = F(b)$

We can repeat this logic for all points.

$\square$

<u>Remark:</u> Here, B_i need not lie in E. However, one can prove that

$$m\left(E \setminus \bigcup_{i=1}^{\tilde{N}} B_i\right) \le 2\delta$$

Given this, we can prove:

<u>Theorem 3.11:</u> (Fundamental Theorem of Calculus?)

Suppose F is absolutely continuous on $[a,b]$. Then,

(i) F' exists a.e. (ii) $F' \in L^1$ (iii) $F(x) - F(a) = \int_a^x F'(t)\,dt$

Conversely, if $f \in L^1([a,b])$, then $\exists$ an absolutely continuous function F with $F' = f$ $\left(\text{in fact, we can take } F(x) = \int_a^x f(t)\,dt\right)$

<u>Proof:</u> We have already seen (i) and (ii). Consider $G(x) = \int_a^x F'(t)\,dt$. We know G is absolutely continuous, and moreover that

$G' = F'$ a.e. by Lebesgue differentiation.

So, $G - F$ is absolutely continuous and $(G-F)' = 0$ a.e..

By Theorem 3.8, $G - F = \text{constant} \Rightarrow G = F + C \underset{\text{set } x=a}{\Longrightarrow} 0 = F(a) + C \Rightarrow C = -F(a)$

$$\Rightarrow \int_a^x F'(t)\,dt = F(x) - F(a)$$

$\square$

<u>Differentiability of Jump Functions</u>

So far we have shown that <u>continuous</u> increasing, bounded functions are differentiable a.e.. We want to remove the continuity assumption.

Note that an increasing, bounded F has at most countably many discontinuities since every jump has a distinct rational in the y-value. Write $\{x_n\}_{n=1}^{\infty}$ for them.

If F has a discontinuity at x_n, set $F(x_n^-) = \lim_{y \uparrow x_n} F(y)$ and $F(x_n^+) = \lim_{y \downarrow x_n} F(y)$

The jump is then $\alpha_n = F(x_n^+) - F(x_n^-)$.

We also have $F(x_n) = F(x_n^-) + \theta_n \alpha_n$ for some $\theta_n \in [0,1]$. Define

$$j_n(x) = \begin{cases} 1 & x > x_n \\ 0 & x < x_n \\ \theta_n & x = x_n \end{cases}.$$

The jump function of F is then $J_F(x) := \sum_{n=1}^{\infty} \alpha_n j_n(x)$

F bounded $\Rightarrow \sum_{n=1}^{\infty} \alpha_n \leq F(b) - F(a) < \infty \Rightarrow \sum_{n=1}^{\infty} \alpha_n j_n(x)$ converges absolutely and uniformly

as a pointwise series

as a sequence of partial sum functions

As continuity is preserved by uniform limits and $j_n(x)$ continuous away from x_n, then J_F is continuous on $[a,b] \setminus \{x_n\}_{n=1}^{\infty}$.

The main lemma is that $F - J_F$ is continuous and increasing. It's also bounded. So, we know $F - J_F$ is differentiable a.e.. To show F is differentiable a.e., it suffices to show J_F is differentiable a.e.

If F has finitely many discontinuities this is obvious. For the infinite case, we use a sort of covering lemma where since $\sum_n \alpha_n < b-a$, most α_n's will be small and we can reduce to the finite case.

Remark: In measure theory, $\mu \ll \nu \iff \nu(A) = 0 \Rightarrow \mu(A) = 0$
For a given ν and any $f \in L^1(\nu)$, we can define a measure $\mu(A) := \int_A f \, d\nu$.

In a sense, this is our characterization of absolute continuity in 1D, where $F(b) = \int_a^b F' \, dx + C$.

So, in 1D we have F abs. cont. $\iff$ $F(b) - F(a)$ abs. cont w.r.t. Lebesgue measure

§4: Hilbert Spaces

Hilbert spaces are crucial because

- they are generalizations of finite-dim spaces to infinite-dim maintaining some rich structure such as orthogonality and angles
- they allow for the framework of analysis to be applied (e.g. infinite sums)

Def:
A Hilbert space $\mathcal{H}$ or $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ is a complete complex inner product space.
It has the following properties:

① $\mathcal{H}$ is a vector space over $\mathbb{C}$ (or $\mathbb{R}$)

② $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ is an inner product:
- $f \mapsto \langle f, g \rangle$ is a linear functional on $\mathcal{H}$ $\forall$ fixed $g \in \mathcal{H}$
- $\langle f, g \rangle = \overline{\langle g, f \rangle}$
- $\langle f, f \rangle \geq 0$ with equality iff $f \equiv 0$

Write $\|f\| := \sqrt{\langle f, f \rangle}$ for the corresponding norm.

③ $\mathcal{H}$ is complete w.r.t. the metric $d(f,g) := \|f - g\|$

Remarks:

① One can prove that Cauchy-Schwartz inequality holds:
$$|\langle f, g \rangle| \leq \|f\| \cdot \|g\|$$

② C.S. $\Rightarrow$ $\|f+g\| \leq \|f\| + \|g\|$ $\Rightarrow$ $\|\cdot\|$ is indeed a norm

③ we will only look at separable $\mathcal{H}$ (i.e. has a countable dense subset)

Examples

① $\mathbb{C}^n$ is a Hilbert space with the usual $\langle (z_1, \dots, z_n), (w_1, \dots, w_n) \rangle = \sum_i z_i \overline{w_i}$
Same with $\mathbb{R}^n$

② $L^2(E) := \{ f \text{ meas, supported on } E, \int_E |f(x)|^2 dx < \infty \}$
with $\langle f, g \rangle = \int_E f \overline{g}$, $E \subseteq \mathbb{R}^n$ with $m(E) > 0$

③ $\ell^2(\mathbb{N}) := \{ (a_n)_{n \in \mathbb{N}} : a_n \in \mathbb{C}, \sum_{n=1}^{\infty} |a_n|^2 < \infty \}$
with $\langle (a_n), (b_n) \rangle = \sum_{n=1}^{\infty} a_n \overline{b_n}$

(4) (a non-separable $\mathcal{H}$)

countable sum!

$$\ell^2(\mathbb{R}) := \left\{ f : \mathbb{C} \to \mathbb{C} \text{ s.t. } \{f \neq 0\} \text{ is countable and } \sum_{x \in \mathbb{R}} |f(x)|^2 < \infty \right\}$$

$$\langle f, g \rangle = \sum_{x \in \mathbb{R}} f(x)\overline{g(x)}, \qquad \|f\| = \left(\sum_{x \in \mathbb{R}} |f(x)|^2 \right)^{\frac{1}{2}}$$

Constructing a Hilbert Space

Def. A semi-inner product is a relation $\langle \cdot, \cdot \rangle$ with the properties

(i) $\forall g$, $f \mapsto \langle f, g \rangle$ is linear (ii) $\langle f, g \rangle = \overline{\langle g, f \rangle}$ (iii) $\langle f, f \rangle \geq 0$

This is the same as an inner product except $\langle f, f \rangle = 0 \not\Rightarrow f = 0$ (i.e. degeneracy)

We can can construct a Hilbert space from such a relation as follows:

(1) Start with a vector space V and a semi-inner product $\langle \cdot, \cdot \rangle$

(2) Define $N := \{ f \in V : \langle f, f \rangle = 0 \}$. Then, $N \subseteq V$ is a linear subspace

(3) Define $\mathcal{H}_0 := V/N$ = equivalence classes of V under $f \sim g \Leftrightarrow f - g \in N$
Note that we can define an inner product on $\mathcal{H}_0$ by

$$\langle f, g \rangle_{\mathcal{H}_0} := \langle f + N, g + N \rangle \Rightarrow \langle f + N, f + N \rangle = 0 \Leftrightarrow \underset{[f] = 0}{f \in N}$$

So, $\mathcal{H}_0$ satisfies (1) and (2). It might not be complete, however.
We call such an $\mathcal{H}_0$ a pre-Hilbert space.

an example of a pre-Hilbert space is the space $\mathcal{H}_0 = \mathcal{R}$ of Riemann-integrable functions on $[-\pi, \pi]$ with the usual inner product

(4) Make $\mathcal{H}_0$ complete.

Prop. 2.7: (make it complete)

Given $(\mathcal{H}_0, \langle \cdot, \cdot \rangle)$ a pre-Hilbert space, we can find a Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ s.t.

(i) $\mathcal{H}_0 \subseteq \mathcal{H}$
(ii) $\langle f, g \rangle_{\mathcal{H}} = \langle f, g \rangle_{\mathcal{H}_0}$ if $f, g \in \mathcal{H}_0$
(iii) $\mathcal{H}_0$ is dense in $\mathcal{H}$

Furthermore, this extension is unique up to isomorphism.
We call this $\mathcal{H}$ the completion of $\mathcal{H}_0$.

"Proof": Consider all Cauchy sequences $\{f_n\}_n \subseteq \mathcal{H}_0$. Define an equivalence relation $\{f_n\}_n \sim \{f_n'\}_n \Leftrightarrow f_n - f_n' \to 0$. Let $\mathcal{H}$ be the equiv. classes. $\square$

Lecture 3/12 -

Last time, we saw pre-Hilbert spaces and orthogonality.

Remark: From the previous proof, we saw Bessel's Inequality:

- $\{e_n\}_n$ orthonormal $\Rightarrow$ Bessel's inequality $\|f\|^2 \geq \sum_i |\langle f, e_i\rangle|^2$
- $\{e_n\}_n$ orthonormal <u>basis</u> $\Rightarrow$ Parseval's identity $\|f\|^2 = \sum_i |\langle f, e_i\rangle|^2$

Theorem 2.4 -

Every separable Hilbert space $\mathcal{H}$ has a countable orthonormal basis.

Proof: $\mathcal{H}$ separable $\Rightarrow \exists$ a countable subset $\{h_k\}_{k=1}^{\infty}$ that is dense
$\Rightarrow \text{span}(\{h_k\}_k)$ is dense in $\mathcal{H}$

WOLOG assume $h_1 \neq 0$. Then, inductively form a new subset $\{\tilde{h}_k\}_k$ as follows:

- $\tilde{h}_1 = h_1$
- if $h_{k+1} \notin \text{span}(\{h_1, \ldots, h_k\})$, include h_{k+1} as the next element in $\{\tilde{h}_k\}_k$

Note that $\text{span}(\{\tilde{h}_k\}_k) = \text{span}(\{h_k\}_k)$ since the elements we were throwing away were already in the span. Also, by construction, $\{\tilde{h}_k\}_k$ is linearly independent. Running Gram-Schmidt (iteratively normalize and subtract parallel components), we get orthonormal $\{f_k\}_k$ which are orthonormal with

$$\text{span}(\{f_k\}_k) = \text{span}(\{\tilde{h}_k\}_k) \Rightarrow \overline{\text{span}(\{f_k\}_k)} = \mathcal{H}$$

saying every $h \in \mathcal{H}$ is an infinite sum of f_k's

$\square$

Remark: If $\mathcal{H}$ has a finite ONB, we say $\mathcal{H}$ is finite-dimensional. Otherwise, it is infinite-dimensional.

Unitary Mappings

Def: We call $U : \mathcal{H}_1 \to \mathcal{H}_2$ between two Hilbert spaces a unitary mapping if

(i) U is linear (ii) U is bijective (iii) $\|U(f)\|_{\mathcal{H}_2} = \|f\|_{\mathcal{H}_1}$ $\forall f \in \mathcal{H}_1$

In other words, U preserves inner products.

Remarks: ① U unitary $\Rightarrow$ U^{-1} is linear bijection and

$$\|U^{-1}(f)\|_{\mathcal{H}_1} = \|U(U^{-1}(f))\|_{\mathcal{H}_2} = \|f\|_{\mathcal{H}_2} \quad \forall f \in \mathcal{H}_2$$

$\Rightarrow U^{-1}$ unitary

② We get that U is 1-Lipschitz (and therefore uniformly continuous) (and an isometry) since (i) and (iii) imply

$$\|U(f)-U(g)\|_{\mathcal{H}_2} = \|U(f-g)\|_{\mathcal{H}_2} = \|f-g\|_{\mathcal{H}_1}$$

③ U unitary $\Rightarrow$ $\langle U(f), U(g)\rangle_{\mathcal{H}_2} = \langle f,g\rangle_{\mathcal{H}_1} \quad \forall f,g \in \mathcal{H}_1$

This follows from the Polarization identity

$$\langle F,G\rangle = \frac{1}{4}\left[\|F+G\|^2 + \|F-G\|^2 + i\left(\|G-iF\|^2 + \|G+iF\|^2\right)\right]$$

So, the inner product is induced by the norm (this happens iff $\|\cdot\|$ satisfies Parallelogram Law $\|F+G\|^2 + \|F-G\|^2 = 2(\|F\|^2 + \|G\|^2)$)

Def: Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$ are unitarily equivalent or unitarily isomorphic if $\exists$ a unitary map $U: \mathcal{H}_1 \to \mathcal{H}_2$. Clearly, this is an <u>equivalence relation.</u>

In fact, all separable Hilbert spaces which are infinite-dimensional are unitarily equivalent to $\ell^2(\mathbb{N})$.

Corollary 2.5

Any two infinite-dimensional separable Hilbert spaces are unitarily equivalent to each other (and to $\ell^2(\mathbb{N})$).

Proof: Fix $\mathcal{H}_1, \mathcal{H}_2$ such Hilbert spaces. Pick ONB's $\{e_k\}_k \subseteq \mathcal{H}_1$, $\{f_k\}_k \subseteq \mathcal{H}_2$. If $f \in \mathcal{H}_1$, then $f = \sum_{k=1}^{\infty} a_k e_k$. Define $U: \mathcal{H}_1 \to \mathcal{H}_2$ by $U(f) = \sum_{k=1}^{\infty} a_k f_k$

By the previous result, U is a bijection. Clearly, U is linear. Also,

$$\|U(f)\|_{\mathcal{H}_2} = \left\|\sum_{k=1}^{\infty} a_k f_k\right\|_{\mathcal{H}_2} = \sum_{k=1}^{\infty} |a_k|^2 = \left\|\sum_{k=1}^{\infty} a_k e_k\right\|_{\mathcal{H}_1} = \|f\|_{\mathcal{H}_1}$$

Parseval

$\square$

In a sense, we map $e_k \mapsto f_k$ and linearly extend.

§ 4.4 - Closed Subspaces and Orthogonal Projections

Def: A (linear) subspace $S \subseteq \mathcal{H}$ is a subset which itself is a vector space.
(i.e. $f, g \in S$, $\alpha, \beta \in \mathbb{C} \Rightarrow \alpha f + \beta g \in S$)

Ex/
① lines through origin in $\mathbb{R}^3$ ② planes through origin in $\mathbb{R}^3$
③ $\{$eventually always 0 sequences$\} \subseteq \ell^2(\mathbb{N})$

Def: A closed subspace $S \subseteq \mathcal{H}$ is a subspace which is closed.
(i.e. $(f_k)_k \subseteq S$ and $f_k \to f \in \mathcal{H} \Rightarrow f \in S$)

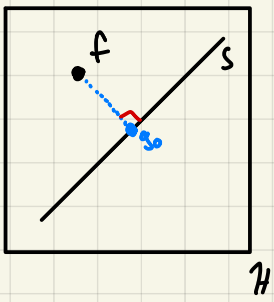
Every finite-dimensional subspace is closed, but not always for infinite-dim (consider example ③ with $f_n = (1, \frac{1}{2}, \dots, \frac{1}{n}, 0, \dots) \in S$, but $f_n \to (1, \frac{1}{2}, \dots) \notin S$)

Also, every closed subspace of a Hilbert space is also a Hilbert space with the induced inner product. Separability is also inherited (see PSet 7)

The crucial property of closed subspaces is that they have (nearest-point) projection maps.

Lemma 4.1: (Existence of orthogonal projection)

Let S be a closed subspace of a Hilbert space $\mathcal{H}$.
Then for any $f \in \mathcal{H}$:



there is a closest point in S (i) $\exists g_0 \in S$ s.t. $\inf_{g \in S} \|f - g\| = \|f - g_0\|$

$f-g_0$ is orthogonal to S, i.e. $f-g_0 \perp S$ (ii) $\forall g \in S$, $f - g_0 \perp g$, i.e. $\langle f - g_0, g \rangle = 0$

Furthermore, g_0 is unique for each $f \in \mathcal{H}$.
We call this g_0 the (orthogonal) projection of f onto S.

We can define the projection map $P_S : \mathcal{H} \to S$ by $P_S(f) := g_0$.

Proof: (i): Set $d := \inf_{g \in S} \|f - g\|$. By definition of inf, we can find $(g_n)_n \subseteq S$ s.t. $\|f - g_n\| \to d$

We want to show $(g_n)_n$ is a Cauchy sequence to show that it converges.
The parallelogram law says: $\|a+b\|^2 + \|a-b\|^2 = 2\left[\|a\|^2 + \|b\|^2\right]$

Applying this with $a = f - g_n$, $b = f - g_m$ gives

$$\underbrace{\|2f - (g_n + g_m)\|^2}_{= 4\|f - \underbrace{\frac{g_n+g_m}{2}}_{\frac{g_n+g_m}{2} \in S \Rightarrow \geq \inf_S}\|^2 \geq 4d^2} + \|g_n - g_m\|^2 = 2\left[\|f-g_n\|^2 + \|f-g_m\|^2\right]$$

$$\Rightarrow \|g_n - g_m\|^2 \leq \underbrace{2\left[\underbrace{\|f-g_n\|^2}_{\to d^2} + \underbrace{\|f-g_m\|^2}_{\to d^2}\right]}_{\to 4d^2} - 4d^2 \Rightarrow \|g_n - g_m\|^2 \xrightarrow[n,m\to\infty]{} 0$$

So, $(g_n)_n$ Cauchy $\Rightarrow$ $g_n \to g_*$ for some $g_* \in \mathcal{H}$ since $\mathcal{H}$ is complete. S closed grants that $g_* \in S$. By continuity of the norm, $\|f - g_n\| \to \|f - g_*\|$. However, by construction, $\|f - g_n\| \to d$. Uniqueness of limits gives $\|f - g_*\| = d$.

(ii) $\forall \varepsilon$ nonzero, consider $g_* - \varepsilon g$ for $g \in S$. As $g_* - \varepsilon g \in S$, we know

$$\|f - (g_* - \varepsilon g)\| \geq \inf_{h \in S} \|f - h\| = \|f - g_*\| \Rightarrow \|(f - g_*) + \varepsilon g\|^2 \geq \|f - g_*\|^2$$

$$\overset{\text{expand w/ inner product}}{\Longrightarrow} \|f - g_*\|^2 + \varepsilon^2\|g\|^2 + 2\varepsilon \operatorname{Re}\langle f - g_*, g\rangle \Rightarrow \varepsilon^2\|g\|^2 + 2\varepsilon \operatorname{Re}\langle f - g_*, g\rangle \geq 0$$

If $\operatorname{Re}\langle f - g_*, g\rangle$ were positive, take ε very small and negative; if $\operatorname{Re}\langle f - g_*, g\rangle$ were negative, take ε very small and positive; either way we can get a contradiction. So, in order for this to hold $\forall \varepsilon$, $\operatorname{Re}\langle f - g_*, g\rangle = 0$.
Repeat this argument with $g_* - i\varepsilon g$ to get $\operatorname{Im}\langle f - g_*, g\rangle = 0$

So, $\langle f - g_*, g\rangle = 0$.

For uniqueness, suppose we have $g_*, \tilde{g}_* \in S$ with $\|f - g_*\| = \|f - \tilde{g}_*\| = d$.
We would then have $g_* - \tilde{g}_* \in S \overset{(ii)}{\Longrightarrow} f - g_* \perp g_* - \tilde{g}_*$

By Pythagoras, $\underbrace{\|f - \tilde{g}_*\|^2}_{d^2} = \|(f - g_*) + (g_* - \tilde{g}_*)\|^2 = \underbrace{\|f - g_*\|^2}_{d^2} + \|g_* - \tilde{g}_*\|^2$

$$\Rightarrow \|g_* - \tilde{g}_*\| = 0$$

$$\Rightarrow g_* = \tilde{g}_*$$

$\square$

Def: For $S \subseteq \mathcal{H}$ subspace, define its orthogonal complement of S by

$$S^\perp := \{ f \in \mathcal{H} : \langle f, g\rangle = 0 \quad \forall g \in S\}$$

Notes:

- $S^\perp$ is always a subspace
- $S^\perp$ is <u>always</u> closed (even if S isn't)

  Cauchy-Schwartz gives $\langle \cdot, g\rangle$ is continuous (in fact, it's $\|g\|$-Lipschitz!) $\forall g \in \mathcal{H}$, since

  $f_n \to f \Rightarrow |\langle f_n, g\rangle - \langle f, g\rangle| = |\langle f_n - f, g\rangle| \leq \underbrace{\|f_n - f\|}_{\to 0} \cdot \|g\|$

  $\Rightarrow |\langle f_n, g\rangle - \langle f, g\rangle| \to 0$. So, $(f_n)_n \subseteq S^\perp$ and $f_n \to f$, then $\forall g \in S$

  $\langle f, g\rangle = \lim_{n\to\infty} \underbrace{\langle f_n, g\rangle}_{=0 \ \forall n} = 0 \Rightarrow f \in S^\perp$

- $S \cap S^\perp = \{0\}$

  $f \in S \cap S^\perp \Rightarrow \langle f, f\rangle = 0 \Rightarrow \|f\|^2 = 0 \Rightarrow f = 0$

Prop 4.2:

If S is a <u>closed</u> subspace of $\mathcal{H}$ Hilbert space, then
$$\mathcal{H} = S \oplus S^{\perp}$$
direct sum: all $f \in \mathcal{H}$ can be written uniquely as $f = f_s + f_{s^\perp}$, $f_s \in S$, $f_{s^\perp} \in S^\perp$

Proof: (Existence) For any $f \in \mathcal{H}$, $\exists g_*$ as in Lemma 4.1, and $f = \underset{\in S}{g_*} + \underset{\in S^\perp \text{ by Lemma 4.1(ii)}}{(f - g_*)}$

(Uniqueness) If $f = a_1 + a_2 = b_1 + b_2$ with $a_1, b_1 \in S$, $a_2, b_2 \in S^\perp$ $\Rightarrow$ $\underbrace{a_1 - b_1}_{\in S} = \underbrace{b_2 - a_2}_{\in S^\perp}$
$\Rightarrow a_1 - b_1, b_2 - a_2 \in S \cap S^\perp \Rightarrow a_1 = b_1, a_2 = b_2$.

$\square$

Def: For a closed subspace S, we can define the **orthogonal projection of $\mathcal{H}$ onto S** $P_s : \mathcal{H} \to S$ by
$$P_s(f) := f_s, \quad \text{where} \quad f = \underbrace{f_s}_{\in S} + \underbrace{f_{s^\perp}}_{\in S^\perp}$$

Properties:
- P_s is linear
- $P_s(f) = f \quad \forall f \in S$
- $P_s(f) = 0 \quad \forall f \in S^\perp$
- $\|P_s(f)\| \leq \|f\| \Rightarrow P_s$ continuous $(\|P_s(f) - P_s(g)\| = \|P_s(f-g)\| \leq \|f - g\|)$

Remark: If $\{e_k\}_k$ is any orthonormal set, then the orthogonal projection onto $\overline{\text{span}\{e_k\}_k}$ is $P(f) = \sum_k \langle f, e_k \rangle e_k$

§4.5: Linear Transformations

Def: If $\mathcal{H}_1, \mathcal{H}_2$ are Hilbert spaces, a **linear transformation** $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a function obeying
$$T(\alpha f + \beta g) = \alpha T(f) + \beta T(g) \quad \forall f, g \in \mathcal{H}, \quad \alpha, \beta \in \mathbb{C}$$
We also call T a **(linear) operator**.

T is **bounded** if $\exists M \geq 0$ s.t. $\|T(v)\| \leq M\|v\| \quad \forall v \in \mathcal{H}_1$
(i.e. $\|T(\frac{v}{\|v\|})\| \leq M \Leftrightarrow \|T(v)\| \leq M \quad \forall \|v\| = 1$)

Def: The **(operator) norm** of an operator $T : \mathcal{H}_1 \to \mathcal{H}_2$ is the smallest $M \geq 0$ that works, and is denoted
$$\|T\|_{op} = \inf\{M \geq 0 : \|T(v)\| \leq M\|v\| \ \forall v \in \mathcal{H}_1\} = \sup_{\|v\|=1} \|T(v)\|$$

We know a linear operator T is continuous iff it's continuous at 0 via shifting. But also:

Lemma:

If T is a linear operator, then

$$T \text{ continuous} \iff T \text{ bounded}$$

Proof: $(\Leftarrow)$ If T bounded, then if $v_n \to v$,

$$\|T(v_n) - T(v)\| = \|T(v_n - v)\| \leq M\|v_n - v\| \to 0$$

$(\Rightarrow)$ Suppose BWOC T is unbounded. So, $\forall M > 0$, $\exists v$ s.t.

$$\|T(v)\| > M\|v\|$$

Take $M > n \in \mathbb{N}$ and get a sequence $(v_n)_n \subseteq \mathcal{H}$ s.t. $\|T(v_n)\| > n\|v_n\|$

$$\Rightarrow \left\|T\left(\frac{v_n}{n\|v_n\|}\right)\right\| > 1. \quad \text{Set } w_n := \frac{v_n}{n\|v_n\|} \Rightarrow \|w_n\| = \frac{1}{n} \Rightarrow w_n \to 0$$

T continuous $\Rightarrow$ $T(w_n) \to T(0) = 0$ $\Rightarrow$ $\|T(w_n)\| \to \|0\| = 0$.

However, $\|T(w_n)\| > 1$ $\forall n$. $\divideontimes$

$\square$

perhaps unnecessary in def.

Def: A **linear functional** ℓ is a continuous linear operator

$$\ell : \mathcal{H} \to \mathbb{C}$$

(forms the dual space)

Ex: $\cdot$ $\forall f \in \mathcal{H}$, the map $\langle \cdot, f \rangle$ is a linear functional

A $\boxed{\text{very special}}$ property of Hilbert spaces is that <u>all</u> linear functionals are of this form. This is the Riesz Representation Theorem.

Lecture 4/19 -

★ **Theorem:** (Riesz Representation)

If ℓ is a continuous linear functional on $\mathcal{H}$ (Hilbert space), then $\exists$ a unique $g \in \mathcal{H}$ s.t.

$$\ell(f) = \langle f, g \rangle \quad \forall f \in \mathcal{H}$$

(i.e. $\mathcal{H}$ is isomorphic to its dual space)

Proof: Let $S := \{f : \ell(f) = 0\}$ be the kernel of ℓ. Note that S is a subspace since ℓ is linear. Furthermore, S is closed because ℓ is continuous. Therefore, $\mathcal{H} = S \oplus S^\perp$. If $\ell = 0$, take $g = 0$ and we are done. So, suppose $S \neq \mathcal{H}$. Then, $S^\perp$ is nonempty. We WTS $S^\perp$ is 1-dimensional.

If $f, g \in S^\perp \setminus \{0\}$ (and so $\ell(f), \ell(g) \neq 0$), then we can write $u = \ell(f) g - f \ell(g)$. Then, $\ell(u) = \ell(f)\ell(g) - \ell(f)\ell(g) = 0 \Rightarrow u \in S$. However, $u \in \text{span}\{f, g\} \Rightarrow u \in S^\perp$.

Thus, $u \in S \cap S^\perp \Rightarrow u = 0 \Rightarrow f, g$ are linearly dependent $\Rightarrow S^\perp$ is 1-dimensional.

Now, take any $h \in S^\perp$ with $\|h\| = 1$, and take $g := \overline{\ell(h)}\, h$.
$\forall f \in \mathcal{H} = S \oplus S^\perp$, $\quad f = \underbrace{f_S}_{\in S} + \underbrace{f_{S^\perp}}_{\in S^\perp} \Rightarrow \ell(f) = \underbrace{\ell(f_S)}_{=0} + \ell(f_{S^\perp}) = \ell(f_{S^\perp})$

Since $\dim(S^\perp) = 1$, h spans $S^\perp \Rightarrow f_{S^\perp} = \langle f_{S^\perp}, h \rangle h$
$$\Rightarrow \ell(f) = \langle f_{S^\perp}, h \rangle \ell(h) = \langle f_{S^\perp}, \overline{\ell(h)} h \rangle = \langle f_{S^\perp}, g \rangle$$
Since $f_S \perp g \Rightarrow \langle f_S, g \rangle = 0$, then $\ell(f) = \langle f_{S^\perp}, g \rangle + \langle f_S, g \rangle = \langle f, g \rangle$, and we are done.

$\square$

Remark:
- If there were 2 such g's, subtract them and it must be 0. Uniqueness follows.
- If $\ell(f) = \langle f, g \rangle$, then $\|\ell\|_{op} = \|g\|_{\mathcal{H}}$

Motivation:
- Spectral theorem for symmetric (normal) matrices says they have an orthonormal eigenbasis. In finite dims, "symmetric" means $A^\dagger = A$; we need an appropriate notion for infinite-dim. We'd like to replace our definitions with some T^* for operators T s.t. $\quad \langle Tx, y \rangle = \langle x, T^* y \rangle$ and $T^{**} = T$.

  We call this the adjoint of T.

Theorem: (Adjoints exist)

Let $T: \mathcal{H} \to \mathcal{H}$ be a linear operator. Then, $\exists$ a unique $T^*: \mathcal{H} \to \mathcal{H}$ obeying

(i) $\langle T(f), g \rangle = \langle f, T^*(g) \rangle \quad \forall f, g \in \mathcal{H}$

(ii) $\|T^*\|_{op} = \|T\|_{op}$

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

Such T^* is called the adjoint of T. We say T is symmetric or self-adjoint if $T = T^*$.

Proof: For any fixed $g \in \mathcal{H}$, define the continuous linear functional $\ell_g(f) := \langle Tf, g \rangle \quad \forall f \in \mathcal{H}$. Riesz Representation gives a unique

$h_g \in \mathcal{H}$ s.t. $\ell_g(f) = \langle f, h_g \rangle \quad \forall f$. If we define $T^*(g) = h_g$, we get $\langle T(f), g \rangle = \ell_g(f) = \langle f, T^*(g) \rangle$

It's easy to see linearity of T^*. Also, since $\|A\|_{op} = \sup_{\|f\|, \|g\| \le 1} |\langle A(f), g \rangle|$, we get

$$\|T\|_{op} = \sup_{\|f\|, \|g\| \le 1} |\langle T(f), g \rangle| = \sup_{\|f\|, \|g\| \le 1} |\langle f, T^*(g) \rangle| = \sup_{\|f\|, \|g\| \le 1} |\langle T^*(g), f \rangle| = \|T^*\|_{op}$$

Lastly, $\langle (T^*)^*(f), g \rangle = \langle f, T^*(g) \rangle = \overline{\langle T^*(g), f \rangle} = \overline{\langle g, T(f) \rangle} = \langle T(f), g \rangle$

Since $(T^*)^*$ and T agree on all inner products, they are equal.

D

Remarks: ① When T is self-adjoint, you can show $\|T\|_{op} = \sup_{\|f\| \le 1} |\langle T(f), f \rangle|$

② $(ST)^* = T^* S^*$

(cool thing: If $L = \frac{d^2}{dx^2}$, then L is self-adjoint via integration by parts!)

Def: Suppose $(\varphi_k)_{k=1}^\infty$ is an ONB of $\mathcal{H}$. Then an operator $T: \mathcal{H} \to \mathcal{H}$ is said to be diagonalized by $(\varphi_k)_{k=1}^\infty$ if

$$T(\varphi_k) = \lambda_k \varphi_k \quad \text{for some } \lambda_k \in \mathbb{C}, \ \forall k$$

In general, if $\varphi \neq 0$ and $\lambda \in \mathbb{C}$ are s.t. $T(\varphi) = \lambda \varphi$, then φ is an eigenvector and λ the corresponding eigenvalue.

Ex If $\mathcal{H} = L^2(\mathbb{R}^n)$ and we define $T: \mathcal{H} \to \mathcal{H}$ by

$$[T(f)](x) := \int_{\mathbb{R}^n} f(y)\, K(x,y)\, dy$$

then we call T an integral operator and K its kernel.
If $K \in L^2(\mathbb{R}^n \times \mathbb{R}^n)$, then T is bounded and we call T a Hilbert-Schmidt operator.

Compact Operators

For finite-dim sets, compact $\Leftrightarrow$ closed and bounded
In infinite-dim, <u>not true</u>: e.g. $\mathcal{H} = \ell^2(\mathbb{Z})$ and sequence $e_n = (0, \dots, 1, 0, \dots)$ (nth index at the 1)
We have $\|e_n\| = 1$ $\forall n$, but no convergence :(

(Minter mentioned that we can prove finite dim $\Leftrightarrow$ unit sphere is compact w.r.t. $\|\cdot\|$ topology)

Def: For $T: \mathcal{H} \to \mathcal{H}$ linear operator, we say T is a compact operator if $\overline{T(B_1)}$ is compact in $\mathcal{H}$, or equivalently:

whenever $(f_n)_n$ is a bounded sequence in $\mathcal{H}$, the sequence $(T(f_n))_n$ has a convergent subsequence. } technically, this is the condition for "sequentially compact"

Notes: ① T compact $\Rightarrow$ T bounded $\Rightarrow$ T continuous

② identity map is <u>not</u> a compact operator on infinite-dim (separable) Hilbert spaces

③ If $\operatorname{rank}(T) < \infty$, then T is compact
(In a sense, compact operators are the closest we can get to finite-dim objects.
E.g., T compact $\Rightarrow$ $\exists$ a sequence of finite rank operators $(T_n)_n$ with $\|T_n - T\| \to 0$)

Lecture 4/24-

Prop 6.1: (Properties of compact operators)

Suppose $\mathcal{H}$ is a Hilbert space and $T: \mathcal{H} \to \mathcal{H}$ is a bounded operator ($T \in \mathcal{B}(\mathcal{H})$). Then,

(i) if $S: \mathcal{H} \to \mathcal{H}$ compact, then ST, TS are compact.
(ii) if $(T_n)_n$ compact and $T_n \to T$ (i.e. $\|T - T_n\|_{op} \to 0$), then T is compact.
(iii) if T is compact, then $\exists$ a sequence $(T_n)_n$ with each T_n having finite rank and $T_n \to T$
(iv) T compact $\iff T^*$ compact

basically converse

(compact is "as close to finite rank" as we can get, as finite rank operators are dense in compact operators)

Proof: (i) Recall that compact $\Rightarrow$ sequence of bounded vectors has a convergent subsequence. Now, $(f_n)_n$ bounded in $\mathcal{H}$ implies $(Tf_n)_n$ bounded $\underset{S\text{ compact}}{\Longrightarrow}$ $(STf_n)_n$ has convergent subsequence $\Rightarrow ST$ compact. For the other one, since T is continuous, a subsequence $(Sf_n)_n$ will converge after applying T. So, TS is compact.

(ii) Let $(f_n)_n$ be a bounded sequence in $\mathcal{H}$. We want to find a convergent subsequence of $(Tf_n)_n$. We use a diagonalization argument.

- T_1 compact $\Rightarrow \exists$ a convergent subsequence $(T_1 f_n)_{n \in A_1}$ for some $A_1 \subseteq \mathbb{N}$ infinite

- T_2 compact $\Rightarrow \exists$ a convergent subsequence $(T_2 f_n)_{n \in A_2}$ for some $A_2 \subseteq A_1$ infinite. Since $A_2 \subseteq A_1$, $(T_1 f_n)_{n \in A_2}$ converges also.

. . .

Repeating inductively, we get $\mathbb{N} \supseteq A_1 \supseteq A_2 \supseteq \dots$ s.t. $\forall N \in \mathbb{N}$, $(T_1 f_n)_{n \in A_N}, (T_2 f_n)_{n \in A_N}, \dots, (T_N f_n)_{n \in A_N}$ converge

Take the diagonal: if k_n is the n^{th} element of A_n, set $\tilde{f}_n := f_{k_n}$. By construction $(T_k \tilde{f}_n)$ converges $\forall k \geq 1$.

Fix $\varepsilon > 0$. As $(f_n)_n$ is bounded, $\|f_n\| \leq C$. So, the triangle inequality gives

$$\|T\tilde{f}_n - T\tilde{f}_m\| \leq \|T\tilde{f}_n - T_k \tilde{f}_n\| + \|T_k \tilde{f}_n - T_k \tilde{f}_m\| + \|T_k \tilde{f}_m - T\tilde{f}_m\|$$

$$\leq \|T - T_k\|_{op} \cdot \|\tilde{f}_n\| + \|T_k \tilde{f}_n - T_k \tilde{f}_m\| + \|T_k - T\|_{op} \|\tilde{f}_m\|$$

Choose k large enough that $\|T - T_k\|_{op} \leq \varepsilon$. For this k, $\forall n, m$ large we know $\|T_k \tilde{f}_n - T_k \tilde{f}_m\| < \varepsilon$ (as $(T_k \tilde{f}_n)_n$ converges). So, $\forall n, m$ large,

$\|T\tilde{f}_n - T\tilde{f}_m\| \leq \varepsilon \cdot C + \varepsilon + \varepsilon \cdot C \Rightarrow (T\tilde{f}_n)_n$ is Cauchy, and so it converges since $\mathcal{H}$ is a Hilbert space. So, T is compact.

(iii) The idea is to project T onto the first n elements of a basis and take $n \to \infty$. Take $\{e_n\}_{n=1}^{\infty}$ be a basis.

Let $Q_n :=$ orthogonal projection of $\mathcal{H}$ onto $\overline{\text{span}\{e_{n+1}, e_{n+2}, \dots\}}$

this is the tail of a convergent series $\sum_k |a_k|^2 = \|g\| < \infty$

Then, $g = \sum_{k=1}^{\infty} a_k e_k \Rightarrow Q_n g = \sum_{k=n+1}^{\infty} a_k e_k \Rightarrow \|Q_n g\|^2 = \sum_{k=n+1}^{\infty} |a_k|^2 \to 0$ as $n \to \infty$ $\forall g$

Suppose by way of contradiction that $\|Q_n T\|_{op}$ doesn't approach 0 as $n \to \infty$. Then, up to a subsequence we have $\|Q_n T\|_{op} \geq C > 0$

$$\Rightarrow \exists g_n \text{ with } \|g_n\| = 1 \text{ and } \|Q_n T g_n\| \geq C \quad \forall n$$

But, T compact $\Rightarrow T g_{n_k} \longrightarrow g \in \mathcal{H}$ for some subsequence ($(g_n)_n$ is bounded).
But then

$$Q_{n_k} g = Q_{n_k}(T g_{n_k}) + Q_{n_k}(g - T g_{n_k})$$

$\|g - T g_{n_k}\| \to 0$ by definition of g, and $\|Q_{n_k} g\| \to 0$ from earlier.

But $\|Q_{n_k}(T g_{n_k})\| \geq C > 0$. $\rightarrow\!\leftarrow$ So, $\|Q_n T\|_{op} \to 0$.

If $P_n :=$ orthogonal proj onto $\text{span}\{e_1, \dots, e_n\}$, then $P_n + Q_n = I$
(rank n; identity)
and

$$\|P_n T - T\|_{op} = \|(P_n - I)T\|_{op} = \|Q_n T\|_{op} \to 0$$

(iv) If T compact, by (iii), $\|P_n T - T\|_{op} \to 0$.
Since adjoints have the same norm, $\|(P_n T - T)^*\|_{op} \to 0$
So, since $P_n^* = P_n$, $\|\underbrace{T^* P_n}_{\text{compact by def. of } P_n} - T^*\|_{op} \to 0$.

By (ii), since compactness is inherited by limits, T^* compact.
If T^* compact, then $(T^*)^* = T$ is also compact. $\square$

Remark:
- If T is diagonalized w.r.t. some basis $\{e_k\}_{k=1}^{\infty}$ and $T e_k = \lambda_k e_k$ for some $\lambda_k \in \mathbb{C}$, then
$$T \text{ compact} \iff \lambda_k \to 0$$
- Hilbert-Schmidt operators are compact

on to the Spectral Theorem!

Theorem 6.2: (Spectral Theorem)

Suppose $\mathcal{H}$ is a separable Hilbert space and $T: \mathcal{H} \to \mathcal{H}$ is a compact, self-adjoint operator.

we call these "symmetric"

Then, $\exists$ an orthonormal basis $\{e_k\}_k$ of $\mathcal{H}$ consisting of eigenvectors of T.

Moreover, if $T(e_k) = \lambda_k e_k$, then $\lambda_k \in \mathbb{R}$ and $\lambda_k \to 0$.

Conversely, if T is any operator defined on $\{e_k\}_k$ in this way, then T is compact and self-adjoint.

We call $\sigma(T) := \{\lambda_k\}_{k=1}^{\infty}$ the spectrum of T.

Idea: We WTS that if $S = \overline{\text{span}\{\text{eigenvectors}\}}$, then $S = \mathcal{H}$.
Suppose BWOC $S \neq \mathcal{H} \Rightarrow \mathcal{H} = S \oplus S^{\perp}$ with $S^{\perp} \neq \{0\}$.
Then, restricting $T|_{S^\perp}: S^\perp \to S^\perp$, if we can find an eigenvector of T in $S^\perp$, we get a nice contradiction.
So, since $S^\perp$ is itself a Hilbert space, the problem reduces to finding a single eigenvector of a symmetric operator in a Hilbert space.

Let's start with the easy parts.

Lemma 6.3:

Suppose T is bounded and self-adjoint. Then,

(i) $\lambda \in \sigma(T) \Rightarrow \lambda \in \mathbb{R}$

(ii) if $f_1 \neq f_2$ are eigenvectors of T with eigenvalues $\lambda_1 \neq \lambda_2$, then $f_1 \perp f_2$

Proof of Lemma: (i) $\lambda \langle f, f\rangle = \langle \lambda f, f\rangle = \langle Tf, f\rangle = \langle f, Tf\rangle = \langle f, \lambda f\rangle = \bar{\lambda}\langle f, f\rangle$

$\underset{\|f\|\neq 0}{\Longrightarrow} \lambda = \bar{\lambda} \Rightarrow \lambda \in \mathbb{R}$

(ii) $\lambda_1 \langle f_1, f_2\rangle = \langle \lambda_1 f_1, f_2\rangle = \langle Tf_1, f_2\rangle = \langle f_1, Tf_2\rangle = \langle f_1, \lambda_2 f_2\rangle = \lambda_2\langle f_1, f_2\rangle$

$\underset{\lambda_1 \neq \lambda_2}{\Longrightarrow} \langle f_1, f_2\rangle = 0 \Rightarrow f_1 \perp f_2$. $\square$

Remark: λ eigenvalue $\Leftrightarrow$ $Tf = \lambda f$ for some $f \neq 0$ $\Leftrightarrow$ $(\lambda I - T)(f) = 0$ $\Leftrightarrow$ $\underbrace{\ker(\lambda I - T)}_{\text{"eigenspace of } \lambda \text{ for } T\text{"}} \neq \{0\}$

Lemma 6.4:

Suppose T is compact and $\lambda \neq 0$. Then,

$$\dim(\ker(\lambda I - T)) < \infty$$

Moreover, for any $\mu > 0$, the subspace spanned by eigenvectors with eigenvalues $> \mu$ is finite-dimensional.

In particular, if $\lambda_1, \lambda_2, \ldots \in \sigma(T)$, then $\lambda_i \to 0$.

Proof of Lemma: Suppose BWOC that $\dim(\ker(\lambda I - T)) = \infty$. Then, we can take $\{\varphi_k\}_{k=1}^{\infty}$ of orthonormal eigenvectors of T with eigenvalue λ.

$\|\varphi_k\| = 1$, T compact $\Rightarrow \{T\varphi_k\}_k$ has convergent subsequence

But $T\varphi_k = \lambda \varphi_k \Rightarrow \|\lambda\varphi_k - \lambda\varphi_m\| = |\lambda| \, \|\varphi_k - \varphi_m\| = \sqrt{2}\,|\lambda| \not\to 0$. $\divideontimes$

they are orthonormal

Using the above, and the fact that different eigenspaces are orthogonal, we are done.

Lemma 6.5:

Suppose $T \neq 0$ is compact and self-adjoint. Then, at least one of $\|T\|_{op}$ or $-\|T\|_{op}$ is an eigenvalue of T.

Proof of Lemma: We use the fact that for self-adjoint T,

$$\|T\|_{op} = \sup\{ |\langle Tf, f\rangle| : \|f\| = 1\}$$

In particular, $\sup\{\langle Tf, f\rangle : \|f\| = 1\} = \|T\|_{op}$ or $-\|T\|_{op}$
So, WOLOG suppose that $= \|T\|$, otherwise consider $-T$.
By def. of supremum, we can take $(f_n)_n$ with $\|f_n\| = 1$ and $\langle Tf_n, f_n\rangle \to \|T\|_{op} =: \lambda$
T compact $\Rightarrow \exists\ (f_{n_k})_k \subseteq (f_n)_n$ s.t. $Tf_{n_k} \to g \in \mathcal{H}$.
Note that $g \neq 0$, since if $g = 0$, then $Tf_{n_k} \to 0$

$\Rightarrow |\langle Tf_{n_k}, f_{n_k}\rangle| \leq \|Tf_{n_k}\| \cdot \|f_{n_k}\| \to 0 \Rightarrow \|T\|_{op} = 0$, $\divideontimes$

So, $g \neq 0$. We claim that $Tg = \lambda g$. Observe that

$$\begin{aligned}
\|Tf_{n_k} - \lambda f_{n_k}\|^2 &= \|Tf_{n_k}\|^2 - 2\lambda \operatorname{Re}\langle Tf_{n_k}, f_{n_k}\rangle + \lambda^2 \|f_{n_k}\|^2 \\
&\leq \|T\|_{op}^2 \|f_{n_k}\|^2 - 2\lambda \underbrace{\operatorname{Re}\langle Tf_{n_k}, f_{n_k}\rangle}_{\text{real b.c. self-adjoint}} + \lambda^2 \underbrace{\|f_{n_k}\|^2}_{\|\cdot\| = 1} \\
&\leq 2\lambda^2 - 2\lambda\langle Tf_{n_k}, f_{n_k}\rangle \\
&\to 0
\end{aligned}$$

So $Tf_{n_k} - \lambda f_{n_k} \to 0 \Rightarrow \lambda f_{n_k} \to g \Rightarrow T(\lambda f_{n_k}) \to Tg$ by continuity of T.

We can now at last prove the Spectral Theorem!

Proof of Spectral Theorem: Let $S := \overline{\text{span}\{\text{eigenvectors of } T\}}$. By Lemma 6.5,

T has an eigenvector $\Rightarrow S \neq \emptyset$. We WTS $S = \mathcal{H}$. Suppose BWOC that $S \subsetneq \mathcal{H}$. Then, $S \oplus S^\perp = \mathcal{H}$ with $S^\perp \neq 0$ a closed, separable Hilbert space.

Note that if $f \in S$, then $Tf \in S$ as T maps eigenvectors to eigenvectors. If $f \in S^\perp$, then $\forall g \in S$, $\langle Tf, g\rangle = \langle \underset{\in S^\perp}{f}, \underbrace{Tg}_{\in S}\rangle = 0 \Rightarrow Tf \perp g \;\; \forall g \in S \Rightarrow Tf \in S^\perp$. So, T maps $S^\perp$ to $S^\perp$.

Now, consider $T' := T|_{S^\perp} : S^\perp \to S^\perp$. T' is also compact and self-adjoint. Certainly, T' can't be 0 since all elements of $S^\perp$ would be eigenvectors. Since $T' \neq 0$, Lemma 6.5 shows that we have an eigenvector $v \in S^\perp$ of T', which is an eigenvector of T. This eigenvector would have to be in $S^\perp$ and in S, which means $v \in S \cap S^\perp \Rightarrow v = 0$. $\ast$

$\square$

Remark: We claimed in the proof of Lemma 6.5 that for self-adjoint T,

$$\|T\|_{op} = \sup\{|\langle Tf, f\rangle| : \|f\| = 1\} =: M$$

To see this, by Cauchy-Schwarz

$$|\langle Tf, f\rangle| \leq \|Tf\| \cdot \|f\| = \|Tf\| \leq \|T\|_{op}$$

For the other direction,

$$\langle Tf, g\rangle = \frac{1}{4}\begin{bmatrix} \langle T(f+g), f+g\rangle - \langle T(f-g), f-g\rangle \\ +i\langle T(f+ig), f+ig\rangle - i\langle T(f-ig), f-ig\rangle \end{bmatrix}$$

For self-adjoint T, $\langle Th, h\rangle \in \mathbb{R}$. So,

$$\begin{aligned} \mathrm{Re}(\langle Tf, g\rangle) &= \frac{1}{4}\left[\langle T(f+g), f+g\rangle - \langle T(f-g), f-g\rangle\right] \\ &\leq \frac{1}{4}\left[M\|f+g\|^2 + M\|f-g\|^2\right] \\ &= \frac{1}{4}\left[2M\|f\|^2 + 2M\|g\|^2\right] \\ &\underset{\|f\| = \|g\| = 1}{=} M \end{aligned}$$

Lastly, we use a little trick

$$|\langle Tf, g\rangle| = e^{i\theta}\langle Tf, g\rangle = \overbrace{\langle Tf, e^{-i\theta} g\rangle}^{\in \mathbb{R}} = \mathrm{Re}(\langle Tf, e^{-i\theta} g\rangle) \leq M$$

In supremum, $\|T\|_{op} \leq M$.

Fun Stuff w/ PDEs

The simplest PDE would have constant coefficients, and be of the form

$$L(u) = f, \quad \text{where} \quad \sum_{|\alpha| \le n} a_\alpha \frac{\partial^\alpha}{\partial x^\alpha} \leftarrow \text{multi-index notation}$$

$\alpha \in \mathbb{N}^n$
$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$
$|\alpha| = |\alpha_1| + \cdots + |\alpha_n|$

and $a_\alpha \in \mathbb{C}$

Fourier Transform maps Schwartz → Schwartz

all partial derivatives decay exponentially at ∞

If f, u are Schwartz functions, we can take the Fourier Transform and get

$$P(\xi)\hat{u} = \hat{f} \quad \text{for some polynomial} \quad P(\xi) := \sum_{|\alpha| \le n} a_\alpha (2\pi i \xi)^\alpha$$

and then take inverse F.T. (requires complex analysis to take the contour integral).

For more general functions, we get around this to find weak solutions (this is related to Sobolev spaces)

Let us work on $\Omega \subseteq \mathbb{R}^n$ open, and let $C_c^\infty(\Omega)$ be smooth + compact support.
Then,

$$C_c^\infty(\Omega) \subseteq L^2(\Omega) \quad \text{is dense}$$

Also, compact support means $|f| = 0$ at $\pm\infty$, and so boundary terms vanish in integration by parts. So, for $\varphi \in C_c^\infty$, $u, f = L(u) \in L^2$, we get

$$\int \varphi L(u) = \int f \varphi = \langle f, \overline{\varphi} \rangle_{L^2}$$

$$\underset{\text{integration by parts}}{=} \int \underbrace{\left(\sum_\alpha a_\alpha (-1)^{|\alpha|} \frac{\partial^\alpha \varphi}{\partial x^\alpha} \right)}_{:= L^* \varphi} u = \langle L^* \varphi, \overline{u} \rangle = \langle u, L^*(\overline{\varphi}) \rangle$$

We say $u \in L^2$ is a weak solution to the PDE $L(u) = f$ if

$$\langle u, L^*(\varphi) \rangle_{L^2} = \langle f, \varphi \rangle_{L^2(\Omega)} \quad \forall \varphi \in C_c^\infty(\Omega)$$

test functions

Theorem:

Suppose $\Omega \subseteq \mathbb{R}^n$ is bounded and open. Suppose L is a linear PDE with constant coefficients.

Then, $\exists$ bounded operator $K: L^2(\Omega) \to L^2(\Omega)$ s.t. $\forall f \in L^2$, $u := K(f)$ is the unique weak solution to $L(u) = f$.

If we define $\mathcal{H} := C_c^\infty$ with $\langle \varphi, \psi \rangle_{\mathcal{H}} = \langle L^*(\varphi), L^*(\psi) \rangle_{L^2}$,
it's a Pre-Hilbert space! Applying Riesz Representation,

$$\ell(\varphi) = \langle f, \varphi \rangle_{L^2} = (g, \varphi)_{\mathcal{H}} = \langle L^* g, L^* \varphi \rangle_{L^2} \quad \forall \varphi \in \mathcal{H}$$

In order for u to be a weak sol. to $Lu = f$, $u := L^* g$!

Def: A weak derivative v of u is a function s.t. $\forall$ test functions $\varphi \in C_c^\infty$,

$$\int u \varphi' = -\int v \varphi \quad (v \text{ obvs}$$

Def: A Sobolev space $W^{k,p}$ is the space of functions in L^p with k weak derivatives, all of which are in L^p.

Thm: (Rellich Compactness Thm)

Let $W^{1,2}$ be Sobolev space with norm $\|u\|_{W^{1,2}}^2 = \int |u|^2 + |\underset{\text{weak derivative}}{Du}|^2$

If $(u_n)_n \subseteq W^{1,2}$ is bounded, then $\exists$ subsequence $(u_{n_k})_k$ s.t.

$$u_{n_k} \to u \text{ in } L^2$$

for some $u \in W^{1,2}$.