Real Analysis

D. Zack Garza

January 7, 2020

Contents

1 Foreword

This are notes from the Fall 2019 Qualifying Exam course for Real Analysis at the University of Georgia, taught by Professor Neil Lyall. These were typeset live during each class. If you find any mistakes or errors, please let me know! Moreover, these may not always reflect exactly what was said or covered in the course – I've paraphrased some things, rephrased others, left out a few bits of more complicated proofs, and added content here and there.

2 Summary

- Measure and integration theory with relevant examples from Lebesgue integration
- Hilbert spaces (only with regard to L^2),
- L^p spaces and the related Riesz representation theorem.
- Hahn, Jordan and Lebesgue decomposition theorems,
- Radon-Nikodym Theorem
- Fubini's Theorem.

Texts

- Real Analysis, by E. M. Stein and R. Shakarchi
- Real Analysis, by G. B. Folland
- An introduction to measure theory, by Terrence Tao
- Real and Complex Analysis, by W. Rudin

An old course page

2.1 Definitions

- **Convolution**

$$f * g(x) = \int f(x-y)g(y)dy$$

- **Dilation**

$$\phi_t(x) = t^{-n}\phi\left(t^{-1}x\right)$$

- The Fourier Transform (todo)

2.2 Convergence Theorems

- **Monotone Convergence Theorem (MCT)**: If $f_n \in L^+$ and $f_n \nearrow f$ a.e., then

$$\lim \int f_n = \int \lim f_n = \int f \quad \text{i.e.} \int f_n \to \int f.$$

- **Dominated Convergence Theorem (DCT)**: If $f_n \in L^1$ and $f_n \to f$ a.e. with $|f_n| \leq g$ for some $g \in L^1$, then

$$\lim \int f_n = \int \lim f_n = \int f \quad \text{i.e.} \int f_n \to \int f,$$

and more generally,

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

Generalized DCT: can relax $|f_n| < g$ to $|f_n| < g_n \to g \in L^1$.

- **Fatou**: If $f_n \in L^+$, then

$$\int \liminf f_n \leq \liminf \int f_n.$$

2.3 Inequalities and Equalities

- **Reverse Triangle Inequality**

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

- **Chebyshev's Inequality**

$$\mu(\{x : |f(x)| > \alpha\}) \leq \left(\frac{\|f\|_p}{\alpha}\right)^p.$$

- ? Inequality

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

with equality iff $a = b$.

- **Holder's Inequality:** For p, q conjugate exponents,

$$\|fg\|_1 \leq \|f\|_p \|g\|_q, \quad \text{i.e.} \int |fg| \leq \left(\int |f|^p\right)^{\frac{1}{p}} \left(\int |g|^q\right)^{\frac{1}{q}}.$$

- **Cauchy-Schwarz**: Set $p = q = 2$ in Holder's inequality to obtain

$$|\langle f,\ g\rangle| = \|fg\|_1 \leq \|f\|_2 \|g\|_2,$$

with equality $\iff$ $f \neq \lambda g$.

- **Minkowski's Inequality:** For $1 \leq p < \infty$,

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p$$

- **Young's Inequality**:

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1 \implies \|f * g\|_r \leq \|f\|_p \|g\|_q.$$

Useful specific cases:

$$\begin{aligned} \|f * g\|_1 &\leq \|f\|_1 \|g\|_1 \\ \|f * g\|_p &\leq \|f\|_1 \|g\|_p, \\ \|f * g\|_\infty &\leq \|f\|_2 \|g\|_2 \\ \|f * g\|_\infty &\leq \|f\|_p \|g\|_q. \end{aligned}$$

- **Bessel's Inequality:** For $x \in H$ a Hilbert space and $\{e_k\}$ an orthonormal sequence,

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

- **Parseval's identity:** Equality in Bessel's inequality, attained when $\{e_k\}$ is a *basis*, i.e. it is complete.

2.3.1 Other

- **Borel-Cantelli Lemma:** Let $\{E_k\}$ be a countable collection of measurable sets. Then

$$\sum_k m(E_k) < \infty \implies \text{ almost every } x \in \mathbb{R} \text{ is in at most finitely many } E_k.$$

- **Egorov's Theorem** Let $E \subseteq \mathbb{R}^n$ be measurable with $m(E) > 0$ and $\{f_k : E \to \mathbb{R}\}$ be measurable functions such that $f(x) := \lim_{k\to\infty} f_k(x) < \infty$ exists almost everywhere.

  Then $f_k \to f$ *almost uniformly*, i.e.

$$\forall \varepsilon > 0, \; \exists F \subseteq E \text{ closed such that } m(E \setminus F) < \varepsilon \;\text{ and }\; f_k \xrightarrow{u} f \text{ on } F.$$

- Fubini
- Tonelli
- Fubini/Tonelli
- **Riemann-Lebesgue Lemma:**

$$f \in L^1 \implies \hat{f}(\xi) \to 0 \text{ as } |\xi| \to \infty.$$

- **Differentiating under the integral**:

If $\left|\frac{\partial}{\partial t} f(x,t)\right| \leq g(x) \in L^1$, then

$$F(t) = \int f(x,t)\ dt \implies \frac{\partial}{\partial t} F(t) := \lim_{h \to 0} \int \frac{f(x,t+h) - f(x,t)}{h} dx$$
$$= \int \frac{\partial}{\partial t} f(x,t)\ dx.$$

Let $h_k \to 0$ be any sequence and define $f_k = \dfrac{f(x,t+h_k) - f(x,t)}{h_k}$, so $f_k \overset{\text{pointwise}}{\to} \dfrac{\partial}{\partial t} f$.

2.4 Important Comments

Measurability:

Best way to show measurability: use Borel characterization, or show that it's an $H \coprod N$ where $H \in F_\sigma$ and N is null.

Just establish something for Borel sets, then use this characterization to extend it to Lebesgue.

AM-GM Inequality:

$$\sqrt{ab} \leq \frac{a+b}{2}.$$

- For finite measure spaces,

$$1 \leq p < q \leq \infty \implies L^q \subset L^p \quad \text{and } \ell^p \subset \ell^q$$

- $C_0([0,1]) \hookrightarrow L^2([0,1])$ is dense.

Dual Spaces: In general, $(L^p)^\vee \cong L^q$

- For qual, supposed to know the $p = 1$ case, i.e. $(L^1)^\vee \cong L^\infty$
  - For the analogous $p = \infty$ case: $L^1 \subset (L^\infty)^\vee$, since the isometric mapping is always injective, but *never* surjective. So this containment is always proper (requires Hahn-Banach Theorem).
- The $p = 2$ case: Easy by the Riesz Representation for Hilbert spaces.

Fourier Series:

- $\hat{f} = \hat{g} \implies f = g$ almost everywhere.

3 First Discussion

3.1 Definitions

Definition: A set X is F_σ iff

$$X = \bigcup_{i=1}^\infty F_i \quad \text{with each } F_i \text{ closed.}$$

Definition: A set X is G_δ iff

$$X = \bigcap_{i=1}^{\infty} G_i \quad \text{with each } G_i \text{ open.}$$

Definition: A set A is *nowhere dense* iff $(\overline{A})^\circ = \emptyset$ iff for any interval I, there exists a subinterval S such that $S \bigcap A = \emptyset$.

Such a set is not dense in *any* (nonempty) open set.

Fact: If the closure of a subset of $\mathbb{R}$ contains no open intervals, it will be nowhere dense.

Definition: A set A is *meager* or *first category* if it can be written as

$$A = \bigcup_{i \in \mathbb{N}} A_i \quad \text{with each } A_i \text{ nowhere dense}$$

Definition: A set A is *null* if for any ε, there exists a cover of A by countably many intervals of total length less than ε, i.e. there exists $\{I_k\}_{j \in \mathbb{N}}$ such that $A \subseteq \bigcup_{j \in \mathbb{N}} I_j$ and $\sum_{j \in \mathbb{N}} \mu(I_j) < \varepsilon$.

If A is null, we say $\mu(A) = 0$.

Some facts:

- If $f_n \to f$ and each f_n is continuous, then D_f is meager.
- If $f \in \mathcal{R}(a, b)$ and f is bounded, then D_f is null.
- If f is monotone, then D_f is countable.
- If f is monotone and differentiable on (a, b), then D_f is null.

We define the *oscillation of f* as

$$\omega_f(x) \coloneqq \lim_{\delta \to 0^+} \sup_{y, z \in B_\delta(x)} |f(y) - f(z)|$$

3.2 Uniform Convergence

Definition: We say that $f_n \to f$ *converges uniformly on A* if

$$\|f_n - f\|_\infty \coloneqq \sup_{x \in A} |f_n(x) - f(x)| \to 0.$$

Note that this defines a sequence of *numbers* in $\mathbb{R}$.

This means that one can find an n large enough that that for every $x \in A$, we have $|f_n(x) - f(x)| \leq \varepsilon$ for any ε.

3.2.1 Showing Uniform Convergence

Find some M_n, independent of x, such that $|f_n(x) - f(x)| \leq M_n$ where $M_n \to 0$.

3.2.2 Negating Uniform Convergence

Fix ε, let n be arbitrary, and find a bad x (which can depend on n) such that $|f_n(x) - f(x)| \geq \varepsilon$.

Example: $\dfrac{1}{1+nx} \to 0$ pointwise on $(0, \infty)$, which can be seen by fixing x and taking $n \to \infty$.

To see that the convergence is not uniform, choose $x = \dfrac{1}{n}$ and $\varepsilon = \dfrac{1}{2}$. Then

$$\sup_{x>0} \left| \frac{1}{1+nx} - 0 \right| \geq \frac{1}{2} \not\to 0.$$

Here, the problem is at small scales – note that the convergence *is* uniform on $[a, \infty)$ for any $a > 0$.

To see this, note that

$$x > a \implies \frac{1}{x} < \frac{1}{a} \implies \left| \frac{1}{1+nx} \right| \leq \left| \frac{1}{nx} \right| \leq \frac{1}{na} \to 0$$

since a is fixed.

3.2.3 Uniformly Cauchy

Let $C^0(([a,b], \|\cdot\|_\infty))$ be the metric space of continuous functions of $[a,b]$, endowed with the metric

$$d(f,g) = \|f-g\|_\infty = \sup_{x\in[a,b]} |f(x) - g(x)|$$

Proposition: This is a complete metric space, and

$$f_n \to^U f \iff \forall\varepsilon\exists N \;\Big|\; m \geq n \geq N \implies |f_n(x) - f_m(x)| \leq \varepsilon \forall x \in X$$

Proof: $\implies$: Use the triangle inequality.

$\impliedby$: Find a candidate limit f: first fix an x, so that each $f_n(x)$ is just a number.

Now we can consider the sequence $\{f_n(x)\}_{n\in\mathbb{N}}$, which (by assumption) is a Cauchy sequence in $\mathbb{R}$ and thus converges.

So define $f(x) \coloneqq \lim_n f_n(x)$.

Aside: we note that if $a_n < \varepsilon$ for all n and $a_n \to a$, then $a \leq \varepsilon$.

Now take $m \to \infty$, we then have

$$\begin{aligned} |f_n(x) - f_m(x)| < \varepsilon \;\forall x &\implies \lim_{m\to\infty} |f_n(x) - f_m(x)| = |f_n(x) - f(x)| \leq \varepsilon \;\forall x \\ &\implies f_n \to^U f. \end{aligned}$$

Note: $f_n \to^U f$ does not imply that $f_n' \to^U f'$.

Counterexample: Let $f_n(x) = \frac{1}{n}\sin(n^2 x)$, which converges to 0 uniformly, but $f_n'(x) = n\cos(n^2 x)$ does not even converge pointwise.

To make this work,

Theorem: If $f_n' \to^U g$ for some g and *for at least 1 point* x, $f_n(x) \to f(x)$, then $g = \lim f_n'$.

3.2.4 Key Example

Exercise: Let

$$f(x) = \sum_{n=1}^{\infty} \frac{nx^2}{n^3 + x^3}.$$

Does it converge at all, say on $(0, \infty)$?

We can check pointwise convergence by fixing x, say $x = 1$, and noting that

$$x = 1 \implies \left|\frac{nx^2}{n^3 + x^2}\right| \leq \left|\frac{n}{n^3 + 1}\right| \leq \frac{1}{n^2} := M_n,$$

where $\sum M_n < \infty$. To see why it does not converge uniformly, we can let $x = n$. Then,

$$x = n \implies \left|\frac{nx^2}{n^3 + x^2}\right| = \frac{n^3}{2n^3} = \frac{1}{2} \not\to 0,$$

so there is a problem at large values of x.

However, if we restrict attention to $(0, b)$ for some fixed b, we have $x < b$ and so

$$\left|\frac{nx^2}{n^3 + x^2}\right| \leq \frac{nb^2}{n^3 + b^2} \leq b^2\left(\frac{n}{n^3}\right) = b^2\frac{1}{n^2} \to 0.$$

Note that this actually tells us that f is *continuous* on $(0, \infty)$, since if we want continuity at a specific point x, we can take $b > x$.

Since each term is a continuous function of x, and we have uniform convergence, the limit function is the uniform limit of continuous functions on this interval and thus also continuous here.

Checking $x = 0$ separately, we find that f is in fact continuous on $[0, \infty)$.

3.3 Series of Functions

Let f_n be a function of x, then we say $\sum_{n=1}^{\infty} f_n$ converges uniformly to S on A iff the partial sums $s_n = f_1 + f_2 + \cdots$ converges to S uniformly on A.

This equivalently requires that

$$\forall \varepsilon \exists N \;\middle|\; n \geq m \geq N \implies |s_n - s_m| = \left|\sum_{k=m}^{n} f_k(x)\right| \leq \varepsilon \quad \forall x \in A.$$

Showing uniform convergence of a series: **Always use the M-test!!!** I.e. if $|f_n(x)| \leq M_n$, which doesn't depend on x, and $\sum M_n < \infty$, then $\sum f_n$ converges uniformly.

Example: Let $f(x) = \sum \frac{1}{x^2 + n^2}$.

Does it converge at all? Fix $x \in \mathbb{R}$, say $x = 1$, then $\frac{1}{1+n^2} \leq \frac{1}{n^2}$ which is summable. So this converges pointwise.

But since $x^2 > 0$, we generally have $\frac{1}{x^2+n^2} \leq \frac{1}{n^2}$ for any x, so this in fact converges uniformly.

3.3.1 Negating Uniform Convergence for Series

Todo

3.4 Misc

A useful inequality:

$$(1+x)^n = \sum_{k=1}^{n} \binom{n}{k} x^k = 1 + nx + n^2 x \geq 1 + nx + nx^2 > 1 + nx$$

Summary of convergence results:

- Functions $f_n \to^U f$:
  - Showing:
    - M test. Produce a bound $\|f_n - f\|_\infty < M_n$ independent of n where $M_n \to 0$.
  - Negating:
    - Each f_n is continuous but f is not,
    - Let n be arbitrary, then find a bad x (which can depend on n) and ε such that

$$\sup |f_n(x) - f(x)| \geq \varepsilon.$$

- Series of functions $\sum f_n \to^U f$:
  - Showing:
    - M test. Produce a bound $\|f_n\|_\infty < M_n$ where $\sum M_n < \infty$.
  - Negating:
    - Each partial sum is continuous but f is not.
    - $f_n \not\to^U 0$.
    - Find a bad x? Work with the partial sums? (Generally difficult?)

4 Tuesday August 15th

See Folland's Real Analysis, definitely a recommended reference.

Possible first day question: how can we "measure" a subset of $\mathbb{R}$? We'd like bigger sets to have a higher measure, we wouldn't want removing points to increase the measure, etc. This is not quite possible, at least something that works on *all* subsets of $\mathbb{R}$. We'll come back to this in a few lectures.

4.1 Notions of "smallness" in $\mathbb{R}$

Definition: Let E be a set, then E is *countable* if it is in a one-to-one correspondence with $E' \subseteq \mathbb{N}$, which includes $\emptyset, \mathbb{N}$.

Definition: A set E is *meager* (or of *1st category*) if it can be written as a countable union of **nowhere dense** sets.

Exercise: Show that any finite subset of $\mathbb{R}$ is meager.

Intuitively, a set is *nowhere dense* if it is full of holes. Recall that a $X \subseteq Y$ is dense in Y iff the closure of X is all of Y. So we'll make the following definition:

Definition: A set $A \subseteq \mathbb{R}$ is *nowhere dense* if every interval I contains a subinterval $S \subseteq I$ such that $S \subseteq A^c$.

Note that a *finite* union of nowhere dense sets is also nowhere dense, which is why we're giving a name to such a *countably infinite* union.

Example: $\mathbb{Q}$ is an infinite, countable union of nowhere dense sets that is not itself nowhere dense.

Equivalently,

- A^c contains a dense, open set.
- The interior of the closure is empty.

We'd like to say a set is *measure zero* exactly when it can be covered by intervals whose lengths sum to less than ε for any $\varepsilon > 0$.

Definition: E is a *null set* (or has *measure zero*) if $\forall \varepsilon > 0$, there exists a sequence of intervals $\{I_j\}_{j=1}^{\infty}$ such that

$$E \subseteq \bigcup_{j=1}^{\infty} \text{ and } \sum |I_j| < \varepsilon.$$

(A second proof of A):

Exercise: Show that a countable union of null sets is null.

We have several relationships

- Countable $\implies$ meager, but not the converse.
- Countable $\implies$ null, but not the converse.

Exercise: Show that the "middle third" Cantor set is not countable, but is both null and meager. Key point: the Cantor set does not contain any intervals.

Theorem: Every $E \subseteq \mathbb{R}$ can be written as $E = A \coprod B$ where A is null and B is meager.

> This gives some information about how nullity and meagerness interact – in particular, $\mathbb{R}$ itself is neither meager nor null. Idea: if meager $\implies$ null, this theorem allows you to write $\mathbb{R}$ as the union of two null sets. This is bad!

Proof: We can assume $E = \mathbb{R}$. Take an enumeration of the rationals, so $\mathbb{Q} = \{q_j\}_{j=1}^{\infty}$. Around each q_j, put an interval around it of size $1/2^{j+k}$ where we'll allow k to vary, yielding multiple intervals around q_j.

To do this, define

$$I_{j,k} = (q_j - 1/2^{j+k}, q_j + 2^{j+k}).$$

Now let $G_k = \bigcup_j I_{j,k}$. Finally, let $A = \bigcap_k G_k$; we claim that A is null.

Note that $\sum_j |I_{j,k}| = \frac{1}{2^k}$, so just pick k such that $\frac{1}{2^k} < \varepsilon$.

Now we need to show that $A^c := B$ is meager. Note that G_k covers the rationals, and is a countable union of open sets, so it is dense. So G_k is an open and dense set. By one of the equivalent formulations of meagerness, this means that G_k^c is nowhere dense.

But then

$$B = \bigcup_k G_k^c$$

is meager.

4.2 $\mathbb{R}$ is not small

Theorem A: $\mathbb{R}$ is not countable.

Theorem B (Baire Category): $\mathbb{R}$ is not meager.

Theorem C: $\mathbb{R}$ is not null.

Note that theorems B and C imply theorem A. You can also replace $\mathbb{R}$ with any nonempty interval $I = [a, b]$ where $a < b$, which is a strictly stronger statement – if any subset of $\mathbb{R}$ is not countable, then certainly $\mathbb{R}$ isn't, and so on.

Proof of (A): Begin by thinking of $I = [0, 1]$, then every number here has a unique binary expansion. So we are reduced to showing that the set of all Bernoulli sequences (infinite length strings of 0 or 1) is uncountable.

Then you can just apply the usual diagonalization argument by assuming they are countable, constructing the table, and flipping the diagonal bits to produce a sequence differing from every entry. ■

A second proof of (A) Take an interval I, and suppose it is countable so $I = \{x_i\}$. Choose $I_1 \subseteq I$ that avoids x_1, so $x_1 \notin I_1$. Choose $I_2 \subseteq I_1$ avoiding x_2 and so on to produce a nested sequence of closed intervals.

Since $\mathbb{R}$ is complete, the intersection $\bigcap_{n=1}^{\infty} I_n$ is nonempty, so say it contains x.

But then $x \in I_1 \in I$, for example, but $x \neq x_i$ for any i, so $x \notin I$, a contradiction. ■

Proof of (B): Suppose $I = \bigcup_{i=1}^{\infty} A_n$ where each A_n is nowhere dense. We'll again construct a nested sequence of closed sets. Let $I_1 \subseteq I$ be a subinterval that misses all of A_1, so $A_1 \bigcap I_1 = \emptyset$ using the fact that A_1 is nowhere dense.

Repeat the same process, let $I_2 \subset I_1 \setminus A_2$. By the nested interval property, there is some $x \in \bigcap A_i$. ■

Note that we've constructed a meager set here, so this argument shows that the **complement of any meager subset of $\mathbb{R}$ is nonempty**. Setting up this argument in the right way in fact shows that this set is dense! Taking the contrapositive yields the usual statement of Baire's Category Theorem.

4.3 Discontinuities

Consider the Thomae function: it is continuous on $\mathbb{Q}$, but discontinuous on $\mathbb{R} \setminus \mathbb{Q}$. Can this be switched to get some function f that is continuous on $\mathbb{R} \setminus \mathbb{Q}$ and discontinuous on $\mathbb{Q}$?

The answer is no. The set of discontinuities of a function is *always* an F_σ set, and $\mathbb{R} \setminus \mathbb{Q}$ is not an F_σ set. Equivalently, the rationals are not a G_δ set.

Let D_f denote the set of discontinuities of f.

Some facts:

- For The pointwise limit of continuous functions, D_f is meager.
- If f is integrable, D_f is null.
- If f is monotone, D_f is countable.
- There is a continuous nowhere differentiable function:
  - Let
    $$f(x) = \sum_n \frac{\|10^n x\|}{10^n},$$
    and in fact *most* functions are like this.
- If f is continuous and monotone, D_f is null.

Theorem: Let $I = [a, b]$. Then

$$I \subseteq \bigcup_{i=1}^{\infty} I_i \implies |I| \leq \sum_{i=1}^{\infty} |I_i|.$$

Proof: The proof is by induction. Assume $I \subseteq \bigcup_{n}^{N+1} I_n$, where wlog we can assume that $a < a_{N+1} < b \leq b_{N+1}$, then $[a, a_{N+1}] \subset \bigcup_{n=1}^{N} I_n$ so the inductive hypothesis applies.

But then

$$b - a \leq b_{N+1} - a = (b_{N+1} - a_{N+1}) + (a_{N+1} - a) \leq \sum_{n=1}^{N+1} |I_n|.$$

■

Note that this proves that $\mathbb{R}$ is uncountable!

5 Thursday August 22nd

Todo: Find notes for first 15 minutes.

5.1 Intervals Are Not Small

Facts:

- Countable $\implies$ Cantor, all intervals are not countable
- Meager $\implies$ Baire, all intervals are not meager
- Null $\implies$ Borel, all intervals are not null.

Exercise: Verify that f is continuous at x iff $\lim f(x_n) = f(x)$ for every sequence $\{x_n\} \to x$.

5.2 Discontinuities

Definition: If $f : X \to \mathbb{R}$, the *oscillation* of f at $x \in X$ is given as

$$\omega_f(x) = \lim_{\delta \to \infty} \sup_{y \in B_\delta(x)} |f(y) - f(z)|.$$

Exercise: Show that f is continuous at $x \iff \omega_f(x) = 0$.

We can then define points of discontinuity as

$$D_f = \left\{ x \in X \;\middle|\; \omega_f(x) > 0 \right\} = \bigcup_{n=1}^{\infty} \left\{ x \in X \;\middle|\; \omega_f(x) \geq \frac{1}{n} \right\}$$

Exercise: Show that D_f is closed.

Theorem 1: f is monotone $\implies$ D_f is countable.

Hint: we can't cover $\mathbb{R}$ by uncountable many disjoint intervals.

Theorem 2: D_f is always an F_σ set.

$\mathbb{R} - \mathbb{Q}$ is not at F_σ set, i.e. one can not construct a function that is discontinuous on exactly this set.

Theorem 3: f is "1st class" $\implies$ D_f is meager.

f is **first class** if $f(x) = \lim_{n \to \infty} f_n(x)$ pointwise and each f_n is continuous.

Theorem 4 (Lebesgue Criterion): Let $f : [a, b] \to \mathbb{R}$ be bounded, then f is Riemann integrable iff D_f is null.

So the Dirichlet function is not Riemann integrable.

Proof of theorems 1 and 2: **Exercise**.

Proof of Theorem 3

We want to show that D_f is meager. We know it's some countable union of some sets, and it suffices to show that they are nowhere dense.

So let $F_n = \left\{x \;\middle|\; \omega_f(x) \geq 0\right\}$ for some fixed n. Let I be an arbitrary closed interval, we will show that there exists a subinterval $J \subseteq I$ with $J \subseteq F_n^c$.

Consider

$$E_k = \bigcap_{i,j \leq k} \left\{x \;\middle|\; |f_i(x) - f_j(x)| \leq \frac{1}{5n}\right\}$$

> Motivation: this comes from working backwards from 4-5 triangle inequalities that will appear later.

Some observations: E_k is closed by the continuity of the f_i **(good exercise)**.

We also have $E_k \subseteq E_{k+1}$. Moreover, $\bigcup_k E_k = \mathbb{R}$ because the $f_i \to f$ are Cauchy.

We'll now look for an interval entirely contained in the complement. Let $I \subset \mathbb{R}$ be an interval, then write

$$I = \bigcup_k (I \bigcap E_k).$$

Baire tells us that I is not meager, so at least one term appearing in this union is *not* nowhere dense, i.e. there is some k for which $I \bigcap E_k$ is not nowhere dense, i.e. it contains an open interval (it has a nonempty interior, and its already closed, and thus it contains an interval).

So let J be this open interval. We want to show that $J \subseteq F_n^c$. If $x \in J$, then $x \in E_k$ as well, and so

$$|f_i(x) - f_j(x)| \leq 1/5n \text{ for all } i, j \geq k.$$

So let $i \to \infty$, so

$$|f(x) - f_j(x)| \leq 1/5n \text{ for all } j \geq k.$$

Now for any $x \in J$, there exists some interval $I(x) \subseteq J$ depending on x such that $|f(y) - f_k(x)| \leq 2/5n$.

Now rewrite this as

$$|f(x) - f_j(x)| = |f(y) - f_k(y) + f_k(y) - f_k(x)|.$$

This implies that $\omega_f(x) \leq 4/5n$. ■

5.3 Integrability

Proof of Theorem 4:

Suppose that $f : [a, b] \to \mathbb{R}$ is bounded.

Recall that f is Riemann integrable iff

$$\forall\ \varepsilon \exists \text{ a partition } P_\varepsilon = \{a = x_1 \leq x_2 \leq \cdots x_n = b\} \text{ of } [a, b] \\ \text{such that } U(f, P_\varepsilon) - L(f, \varepsilon) \leq \varepsilon,$$

where

$$U(f, P_\varepsilon) - L(f, \varepsilon) \coloneqq \sum_n \sup_{y,z \in [x_n, x_{n+1}]} |f(y) - f(z)|(x_{n+1} - x_n)$$

($\Rightarrow$): Let $\varepsilon > 0$ and n be fixed, and produce a partition P_ε so that this sum is less than ε/n.

Recall that we want to show that F_n is null.

Now exclude from this sum all intervals that miss F_n, making it no bigger. We also know that in F_n, the sups are no greater than $1/n$,

$$\varepsilon/n \geq \sum \text{stuff} \geq \sum \frac{1}{n}(x_{n+1} - x_n)$$

■

($\Leftarrow$): Suppose D_f is null and let $\varepsilon > 0$ be arbitrary, we want to construct P_ε. Choose $n > 1/\varepsilon$ and $F_n \subseteq D_f$ is closed and bounded and thus compact.

But a compact measure zero interval can in fact be covered by *finitely* many open intervals.

So F_n is covered by finitely many intervals $\{I_n\}_{n=1}^N$ such that $\sum |I_n| \leq \varepsilon$.

Now if $x \notin F_n$, then $\exists \delta(x) > 0$ where

$$\sup_{y,z \in B_\delta(x)} |f(y) - f(z)| < \frac{1}{n} < \varepsilon.$$

Since $(\bigcup_j I_j)^c$ is compact, there's a finite cover $I_{N+1}, \cdots I_{N'}$ covering F_n^c.

■

6 Tuesday August 27th

6.1 Nowhere Density

Recall **Baire's Theorem**: $\mathbb{R}$ can not be written as a countable union of nowhere dense sets.

A subset $A \subseteq \mathbb{R}$ is *nowhere dense* $\iff$ every interval contains a subinterval which lies entirely in $A^c \iff \overline{A}$ has empty interior $\iff \overline{A}$ contains no open intervals.

Exercise: Show that these definitions are equivalent.

Corollary: $\mathbb{R} \setminus \mathbb{Q}$ is not an F_σ set.

Proof: Suppose it was, so

$$\mathbb{R} \setminus \mathbb{Q} = \bigcup_{n \in \mathbb{N}} F_n \text{ with } A_n \text{ closed }.$$

Then

$$\mathbb{R} = \left(\bigcup_n F_n \right) \bigcup \left(\bigcup_i \{q_i\} \right) \text{ where } \mathbb{Q} = \bigcup \{q_i\}.$$

This exhibits $\mathbb{R}$ as a countable union of closed sets. But the F_n are nowhere dense, since if they contained in interval they'd also contain a rational. ■

Exercise: Show that F_n are nowhere dense by constructing a sequence of elements in F_n that converges to an element in $F_n^c \subset \mathbb{Q}$.

6.2 Riemann Integration

Some nice properties:

- Good for approximation (vertical strips)
- Many functions are in $\mathcal{R}$, e.g. continuous functions.
- $\mathcal{R}([a, b])$ is a vector space
- The integral is an element of $\mathcal{R}^*$.
- FTC
- $\mathcal{R}$ is closed under uniform convergence.

Some bad properties:

- The Dirichlet function $\mathbb{1}[x \in \mathbb{Q}]$ is not in $\mathcal{R}$. **(Exercise!)**
  - **Exercise**: Show that $D_f = \mathbb{R}$ (use sequential continuity)
  - It is in $\mathcal{L}$ (Lebesgue integral).
- $\mathcal{R}$ is not closed under pointwise convergence.
  - Example: $g_n(x) = \mathbb{1}\left[x \in \mathbb{Q},\ x = \frac{p}{q}, q \leq n\right] \in \mathcal{R}$, but $g_n \not\rightrightarrows g$. **(Exercise)**

In fact, there exists a sequence of *continuous* functions $\{f_n\}$ such that

- $0 \leq f_n(x) \leq 1$ for all x, n.
- $f_n(x)$ is decreasing as $n \to \infty$ for all x.
- $f := \lim f_n \notin \mathcal{R}$.

This seems disturbing! The Lebesgue integral fixes this particular problem. Letting

$$\mathcal{L} = \left\{ f \;\middle|\; f \text{ is Lebesgue integrable} \right\},$$

we have the following theorem:

Theorem (Dominated Convergence, Special Case):

Let $\{f_n : [a, b] \to \mathbb{R}\} \subset \mathcal{L}$, such that

$$\forall n \in \mathbb{N}, \forall x, |f_n(x)| \leq M.$$

If $f_n \to f$ pointwise then $\displaystyle\int f_n \to \int f$.

6.3 Measure Theory

6.3.1 The Non-Measurable Set

Can we assign a "measure" to all subsets of $\mathbb{R}^n$?

This should be a function $m : \mathcal{P}(\mathbb{R}^n) \to \mathbb{R}^{\geq 0} \bigcup \{\infty\} = [0, \infty]$ with some properties (see handout).

- If $\{E_i\}_{i\in\mathbb{N}}$ are disjoint, then $m(\coprod_{i\in\mathbb{N}} E_i) = \displaystyle\sum_{i\in\mathbb{N}} m(E_i)$.
- If $E \simeq F$ by translation/rotation/reflection, then $m(E) = m(F)$.
- $m(Q) = 1$ if $q = [0, 1]^n$.

But so far, this is impossible for the following reason:

Proposition: There exists a non-measurable set:

Proof: Define an equivalence relation $x \sim y \iff x - y \in \mathbb{Q}$ on $[0, 1)$. Note that each equivalence class bijects with $\mathbb{Q}$, so each class is countable and there must be an uncountable number of classes.

Use the axiom of choice to construct a set N by choosing exactly one element from each equivalence class.

Now let $\mathbb{Q} \bigcap [-1, 1] = \{q_j\}$ be an enumeration of the rationals in this interval, and define $N_j = N + q_j$.

Note

$$j \neq k \implies N_j \bigcap N_k = \emptyset.$$

By translation invariance, $m(N_j) = m(N)$, and

$$[0, 1) \subseteq \coprod_j N_j \subseteq [-1, 2].$$

But then by taking measures and using the fact that $m(N_i) = m(N)$, we have

$$1 \leq \sum_j m(N_j) \leq 3,$$

But then $m(N) = 0 \implies 1 > m(N)$, and if $m(N) = \varepsilon > 0$ then

$$\sum m(N) = \sum \varepsilon > 3.,$$

a contradiction. ■

Exercise: Any open set in $\mathbb{R}$ can be written as a *countable* union of intervals.

But what can be said about closed sets, or all of $\mathbb{R}^n$?

Fact: Any open set in $\mathbb{R}^n$ can be written as an *almost disjoint* union of closed cubes.

We can then attempt to ascribe a measure to a set by approximating an open set from the inside by cubes. However, it's not clear that this is unique (although it is).

7 Wednesday August 28th

7.1 Outer Measure

Definition (Lebesgue Outer Measure): For any $E \subseteq \mathbb{R}^n$ define

$$m_*(E) = \inf \sum |Q_i|$$

where the infimum is taken cover all countable coverings of E by closed cubes Q_i.

Proof of property (4): Since we have property (2), we just need to show

$$m_*(E_1 \bigcup E_2) \geq m_*(E_1) + m_*(E_2).$$

Choose δ such that $0 < \delta < \mathrm{dist}(E_1, E_2)$ and let $\varepsilon > 0$.

Then there exists a covering of $E_1 \bigcup E_2$ such that

$$m_*(E_1 \bigcup E_2) \leq \sum |Q_i| \leq m_*(E_1 \bigcup E_2) + \varepsilon.$$

We can assume (possibly after subdividing) that $\mathrm{diam}(Q_i) < \delta$. Then each Q_i can intersect at most one of E_1, E_2.

Let

$$J_1 = \left\{ j \;\middle|\; Q_j \bigcap E_1 \neq \emptyset \right\}$$
$$J_2 = \left\{ j \;\middle|\; Q_j \bigcap E_2 \neq \emptyset \right\}.$$

Note that J_1, J_2 are disjoint, and we have

$$E_1 \subseteq \bigcup_{j \in J_1} Q_j$$
$$E_2 \subseteq \bigcup_{j \in J_2} Q_j.$$

But then

$$m_*(E_1) + M_*(E_2) \leq \sum_{j \in J_1} |Q_j| + \sum_{j \in J_2} |Q_j|$$

by definition, since m_* is an infimum.

But this is less than summing over *all* j, which is the term appearing in the cover we choose above.
■

7.2 Covering by Cubes

Proof of property (5):

Qual/Exam problem alert (DZG)

Let $\varepsilon > 0$, we will show

$$\sum |Q_j| \leq m_*(E) + \varepsilon.$$

Start by shrinking the cubes. Choose

$$\tilde{Q}_j \leq |Q_j| \leq \left|\tilde{Q}_j\right| + \varepsilon/2^j.$$

Then for any finite N, any collection of N different Q_js are disjoint.

Exercise: If K is compact, F is closed, and $K \bigcap F = \emptyset$, the $\text{dist}(K, F) > 0$.

Note that although this is certainly true for the *entire* infinite collection, we take finitely many so we can get a δ that uniformly bounds the distance between any two from below.

But then

$$m_*(\bigcup_{j=1}^N \tilde{Q}_j) = \sum_{j=1}^N \left|\tilde{Q}_j\right| \geq \sum_{j=1}^N \left|\tilde{Q}_j\right| - \varepsilon$$

and since $\displaystyle\bigcup_{j=1}^N \tilde{Q}_j \subseteq E$, for all N we have

$$\sum_{j=1}^N \left|\tilde{Q}_j\right| - \varepsilon \geq m_*(E).$$

(Missing details, finish/fill in.)

Definition (Measurable):

A set $E \subseteq \mathbb{R}^n$ is (Lebesgue) measurable if

$$\forall \varepsilon > 0 \quad \exists G \text{ open} \quad \Big| \quad m_*(G \setminus E) < \varepsilon.$$

Important Observations:

- If $m_*(E) = 0$, the E is automatically measurable.
- If $F \subseteq E$ and $m_*(E) = 0$, then F is automatically measurable.

8 Thursday August 29th: Outer Measure

Today: 1.2 in Stein, the Lebesgue Outer Measure

Some preliminary results about open sets:

Theorem 1.3 (Stein): Every open subset of $\mathbb{R}$ can be written as a countable union of disjoint open intervals. Moreover, this representation is unique.

Theorem 1.4 (Stein): As a partial analog of 1.3, every subset of $\mathbb{R}^n$ can be written as a countable union of *almost disjoint* closed cubes.

Definition: A, B are *almost disjoint* iff $A^\circ \bigcap B^\circ = \emptyset$.

We'll now attempt to assign a preliminary notion of measure for all subsets of $\mathbb{R}$ which extends the notion of volume.

Definition (Outer Measure): If $E \subseteq \mathbb{R}^n$, then the **Lebesgue outer measure** of E, denoted $m_*(E)$, is defined as follows:

$$m_*(E) = \inf_{\bigcup_{i\in\mathbb{N}} Q_i \supseteq E,\ Q_i \text{closed}} \sum_{i\in\mathbb{N}} |Q_j|$$

where we take the infimum over all coverings of E by countably many closed cubes.

Remarks:

- m_* is well-defined for *all* subsets $E \subseteq \mathbb{R}^n$.
- $m_*(E) \in [0, \infty]$
- For all $\varepsilon > 0$, there exists a covering $E \subseteq \bigcup_{i\in\mathbb{N}}$ such that

$$\sum_{i\in\mathbb{N}} |Q_i| \leq m_*(E) + \varepsilon.$$

- We would not want to merely require coverings by *finite* collections of closed cubes. (See challenge problem and *Jordan content* of sets)

Examples: If E is countable, then $m_*(E) = 0$.

This follows from the fact that any point is a closed cube with zero volume

Proposition: If $E \subset \mathbb{R}$, then E is null $\iff m_*(E) = 0$.

$\implies$: We can cover by open intervals with lengths summing to $< \varepsilon$, so just close them (which doesn't increase the length).

$\impliedby$: **(Easy exercise.)**

Increase the length of the nth open interval by $\varepsilon/2^n$. ■

Example: If Q is a closed cube, the $m_*(Q) = |Q|$, the usual volume.

Since $Q \subseteq Q$, Q covers itself and $m_*(Q) \leq |Q|$. For the other direction, fix $\varepsilon > 0$; we will show $|Q| \leq m_*(G) + \varepsilon$ for every ε.

Let $\{Q_j\}$ be an arbitrary covering of Q by closed cubes. Idea: enlarge the cubes a bit so they're open, and use compactness to get a finite subcover.

Let S_j denote an open cube with the property that $Q_j \subseteq S_j$ and

$$|Q_j| \leq |S_j| \leq (1+\varepsilon)|Q_j|$$

Since Q is compact, there is a finite N such that $E \subseteq \bigcup_{j=1}^{N} S_j$, and the claim is that $|Q| \leq \sum_{j=1}^{N} |S_j|$ (Lemma 1.2, Stein).

Recall 1-dimensional setting, we did the same thing to prove that $\mathbb{R}$ was not null.

We then have

$$|Q| \leq \sum_{j=1}^{N} |S_j| \leq (1+\varepsilon) \sum_{j=1}^{N} |Q_j| \leq (1+\varepsilon) \sum_{j=1}^{\infty} |Q_j|,$$

which is what we wanted to show.

Exercise: Let Q be open, show that $m_*(Q) = |Q|$.

Proposition:

$$m_*(\mathbb{R}^n) = \infty.$$

This would follow if we could show that $|Q| \leq m_*(\mathbb{R}^n)$ for any Q, and we can take Q to be arbitrarily large. This is because any covering of $\mathbb{R}^n$ is also a covering of Q.

8.1 Properties of Outer Measure

1. **Monotonicity**: If $E_1 \subseteq E_2$ then $m_*(E_1) \leq m_*(E_2)$.
2. **Countable Subadditivity**: If $E = \bigcup_{i \in \mathbb{N}} E_i$ for any countable union, then

$$m_*(\bigcup E_i) \leq \sum m_*(E_i)$$

3. If $E \subseteq \mathbb{R}^n$ then $\forall \varepsilon > 0$ there exists an *open* set $G \supseteq E$ such that

$$m_*(E) \leq m_*(G) \leq m_*(E) + \varepsilon.$$

Note: This does not imply every set is measurable! In fact,

$$m_*(G) - m_*(E) \neq m(G \setminus E).$$

If we try to write $G = E \coprod (G \setminus E)$, we only get an equality if there's a positive distance between these two sets! Otherwise, we only have subadditivity, and

$$m_*(G \setminus E) \geq m_*(G) - m_*(E).$$

But this is the wrong direction if we want to say something like

$$m_*(G) - m_*(E) \leq \varepsilon.$$

4. **Almost Disjoint Additivity:** If $E_1, E_2 \subseteq \mathbb{R}^n$ and

$$\operatorname{dist}(E_1, E_2) := \inf_{x \in E_1, y \in E_2} |x - y| > 0 \implies$$
$$m_*(E_1 \dot{\bigcup} E_2) = m_*(E_1) + m_*(E_2).$$

5. If $E = \dot{\coprod}_{j \in \mathbb{N}} Q_j$ with the Q_j *almost disjoint*, then $m_*(E) = \sum_j |Q_j|$.

Remark: Property 4 does *not* hold in general if we merely assume that $E_1 \bigcap E_2 = \emptyset$. It *will* be true if we restrict the collection of sets we consider to be "measurable", so any counterexample will have to involve a pathological set.

Warning: Any E_j could have $m_*(E) = \infty$, so we have to be careful with our assumptions and how we work with inequalities, particularly when subtracting measures.

8.2 Proofs of Properties of Outer Measure

8.2.1 Property 1

Straightforward, since any covering of E_2 is also a covering of E_1.

We are thus taking infimums over *larger* collections of sets, so it can only get smaller.

8.2.2 Property 2

If $m_*(E_j) = \infty$ for any j, this is vacuous, so assume $m_*(E_j) < \infty$ for every j.

Let $\varepsilon > 0$. For each j, there exists a covering $E_j \subseteq \bigcup_k Q_{j,k}$ where

$$\sum_k |Q_{j,k}| \leq m_*(E_j) + \varepsilon/2^j.$$

But now $E \subseteq \bigcup_{j,k} Q_{j,k}$, so

$$m_*(E) \leq \sum_{j,k} |Q_{j,k}| = \sum_j \sum_k |Q_{i,j}| \leq \sum_j (m_*(E_j) + \varepsilon/2^j) = \varepsilon + \sum_j m_*(E_j).$$

8.2.3 Property 3

Idea: enlarge open sets in a summable way.

Let $\varepsilon > 0$. Then there exists a covering $E \subseteq \bigcup_{j \in \mathbb{N}} Q_j$ such that

$$\sum |Q_j| \leq m_*(E) + \varepsilon/2.$$

Let S_j be an open cube such that $Q_j \subset S_j$ and

$$|S_j| \leq |Q_j| + \varepsilon/2^{j+1}.$$

So define $G := \bigcup_j S_j$, which is open.

Using subadditivity, we have

$$m_*(G) \leq \sum_j |S_j| = \sum_j \left(|Q_j| + \varepsilon/2^{j+1}\right) \leq m_*(E) + \varepsilon.$$

8.2.4 Property 4

We Just need to show that

$$m_*(E_1 \bigcup E_2) \leq m_*(E_1) + m_*(E_2),$$

since the reverse direction follows from (2).

Proof to follow in next section.

Key idea: by subdividing cubes, we can assume that no cube intersects both sets.

Remark: It is possible to construct closed disjoint subsets of $\mathbb{R}^n$ such that the distance between them is still zero. Take $X = \mathbb{N}$ and $Y = \left\{n + \frac{1}{2n} \;\middle|\; n \in \mathbb{N}\right\}$.

Exercise (a good one!): Show that if F is closed and K is **compact**, then $\mathrm{dist}(X, Y) > 0$.

9 Tuesday September 3rd

9.1 Lebesgue Measurability

Recall the definition of the Lebesgue measure:

Definition: For any $E \subseteq \mathbb{R}^n$, we define

$$\mu(E) = \inf\left\{\sum |Q_i| \;\middle|\; E \subset \bigcup Q_i, Q_i \text{ a closed cube}\right\}$$

This satisfies properties (1) through (5).

Note we don't have finite additivity for the outer measure.

Definition: A set E is said to be *measurable* iff

$$\forall \varepsilon > 0, \ \exists \text{ an open } G \supseteq E \;\Big|\; \quad m_*(G \setminus E) < \varepsilon.$$

Observations:

- If E is open, E is measurable
- If $m_*(E) = 0$, then E is measurable. (Quite a special property!)

- If E is closed, E is measurable. (Needs a proof.)

Theorem 1: The collection $\mathcal{M}$ of all measurable sets is a σ-*algebra*, i.e. $\mathcal{M}$ is closed under

- Countable unions
- Complements
- Countable intersections

Theorem 2: The Lebesgue measure (on measurable sets) is countably additive, i.e. if $\{E_i\}_{i\in\mathbb{N}}$ is a countable collection of *pairwise disjoint* measurable sets, then

$$m(\coprod E_i) = \sum m(E_i).$$

9.2 Lebesgue Measurable Sets Form a Sigma Algebra

Proof of Theorem 1:

Part 1: Let $E = \bigcup_{i\in\mathbb{N}} E_i$ with each E_i measurable; we want to show E is measurable.

Given $\varepsilon > 0$, for each j choose $G_j \supseteq E_j$ such that

$$m_*(G_j \setminus E_j) < \varepsilon/2^j.$$

Then let $G = \bigcup G_j$, which is open and $G \supseteq E$ and $G \setminus E = \bigcup G_j \setminus E_j$.

Using monotonicity, and then subadditivity, we have

$$m_*(G \setminus E) = m_*(\bigcup G_j \setminus E_j) \leq \sum m_*(G_j - E_j) \leq \sum \varepsilon/2 = \varepsilon.$$

Part 2: Let $E \in \mathcal{M}$. Then for all $k \in \mathbb{N}$, there is an open $G_k \supseteq E$ with

$$m_*(G_k \setminus E) \leq \frac{1}{k}.$$

Lemma to prove later: G_k^c is closed and measurable.

By (1), the set $S \coloneqq \bigcup G_k^c$ is measurable, and $S \subseteq E^c$, since $E^c = S \bigcup (E^c \setminus S)$. So we just need to show that $E^c \setminus S = E^c \bigcap S^c$ is measurable.

But where does S live? Since

$$E^c \setminus S \subset G_k \setminus E = G_k \bigcap E^c \text{ for every } k,$$

we have

$$m_*(E^c \setminus S) \leq m_*(G_k \setminus E) < \frac{1}{k} \text{ for all } k,$$

which says that $m_*(E^c \setminus S) = 0$ and thus $E^c \setminus S$ is measurable.

■

Think further about why outer measure zero sets should be Lebesgue measurable!

Next time:

- Closed sets are measurable,
- Proof of theorem 2,
- Characterizations of measurability.

10 Thursday September 5th

10.1 Measurability of Closed Sets

Recall: A set E is *Lebesgue measurable* iff there exists an open set G with $E \subseteq G$ and $m(G \setminus E) < \varepsilon$ for any $\varepsilon > 0$, and the set $\mathcal{M}$ of all measurable sets forms a σ-algebra.

Fact: If $F, K \in \mathbb{R}^n$ with F closed, K compact, and $F \bigcap K = \emptyset$, then $\text{Dist}(F, K) > 0$.

Proof: Towards a contradiction, suppose the distance in zero. Idea: we'll use sequential compactness.

We can produce sequences $\{x_n\} \subset F, \{y_n\} \subset K$ such that $|x_n - y_n| \to 0$.

Since K is compact, it is *sequentially compact*, so there is a subsequence $\{y_{n_k}\}$ with $y_{n_k} \to y \in K$.

Then

$$|x_{n_k} - y| \leq |x_{n_k} - y_{n_k}| + |y_{n_k} - y| \to 0.$$

$\blacksquare$

We used the following:

Lemma: Closed sets are measurable.

Proof:

Claim: It suffices to prove this for *compact* sets.

Let F be closed. Then write $F = \bigcup_k (F \bigcap B(k, 0))$. But $F \bigcap B(k, 0)$ is closed and bounded, thus compact.

So if we show compact sets are measurable, we've written F as a countable union of measurable sets, which is thus measurable.

So suppose K is compact, we want to show that $m_*(K) < \infty$. Given $\varepsilon > 0$, we can find an open set $G \supseteq K$ such that

$$m_*(G) < m_*(K) + \varepsilon.$$

Now, *since K is bounded*, the outer measure is not infinite, and so we have $m_*(G) - m_*(K) < \varepsilon$.

Goal: We now want to show

$$m_*(G \setminus K) \leq m_*(G) + m_*(K).$$

Since G is open, $G \setminus K$ is open as well. We can write any open set as the union of almost disjoint closed cubes, so we have

$$G \setminus K = \bigcup_j Q_j, \quad \{Q_j\} \text{ a collection of almost disjoint cubes.}$$

Now by property (5), we have

$$m_*(G \setminus K) = \sum_j |Q_j|.$$

Since any finite union of closed sets is closed, we have $K \bigcap (\bigcup_{j=1}^{N} Q_j) = \emptyset$. But then $\text{dist}(K, \bigcup^{N} Q_j) > 0$. Using (1) and (4),

$$\begin{aligned}
m_*(G) \geq m_*(K \bigcup_{j=1}^{N} Q_j) &= m(K) \\
&= m(\bigcup^{N} Q_j) \\
&= m_*(K) + \sum^{N} |Q_j|, .
\end{aligned}$$

and since K is bounded and thus of finite measure, we obtain

$$\sum^{\infty} |Q_j| \leq m_*(G) - m_*(K).$$

■

10.2 Characterizations of Measurability

Theorem: A set $E \subseteq \mathbb{R}^n$ is measurable iff

1. For any ε, $\exists F \subseteq E$ with F closed and $m_*(E \setminus F) < \varepsilon$.
2. There exist F closed, G open, $F \subseteq E \subseteq G$ with $m_*(G \setminus F) < \varepsilon$.

We know that E is measurable iff E^c is measurable, so we'll apply the definition to E^c. So we know

$$\forall \varepsilon > 0, \ \exists \text{ open } G \supseteq E^c \ \Big| \ m_*(G \setminus E^c) < \varepsilon.$$

and so

$$\forall \varepsilon > 0, \ \exists \text{ closed } G^c \subseteq E \ \Big| \ m_*(E \setminus G^c) < \varepsilon.$$

since $G \setminus E^c = G \bigcap E = E \setminus G^c$. So just take $F = G^c$ and we're done.

Definition: A σ-algebra is any collection of sets which is closed under complements and countable unions.

Note that if we intersect σ-algebras, we still get a σ-algebra.

Examples:

- $\mathcal{P}(\mathbb{R}^n)$

- $\mathcal{M}$, the collection of all (Lebesgue) measurable sets.
- $\mathcal{B}(\mathbb{R}^n)$, the *Borel* subsets of $\mathbb{R}^n$, i.e. the smallest σ-algebra containing the open sets.

There are inclusions

$$\mathcal{P}(\mathbb{R}^n) \supsetneq \mathcal{M}(\mathbb{R}^n) \supsetneq \mathcal{B}(\mathbb{R}^n).$$

Qual problem alert!

Theorem: TFAE:

1. $E \subseteq \mathbb{R}^n$ is measurable
2. $E = H \bigcup Z$ where $H \in F_\sigma$ and Z is null
3. $E = V \setminus Z'$ where $V \in G_\delta$ and Z' is null.

Proof:

$2, 3 \implies 1$ **Exercise.** This is the easy direction.

$1 \implies 2, 3$: For all $k \in \mathbb{N}$, we can find $F_k \subseteq E \subseteq G_k$ with F_k closed, G_k open, and

$$m_*(G_k \setminus F_k) < \frac{1}{k}.$$

So let $V = \bigcap G_k$ and $H = \bigcup F_k$. Then $H \subseteq E \subseteq V$. Note that H is an F_σ and V is a G_δ. Moreover, $V \setminus H \subseteq G_k \setminus F_k$ for all k. By subadditivity,

$$m_*(V \setminus H) \leq m_*(G_k \setminus F_k) \to 0.$$

Now, $E = H \bigcup (E \setminus H)$ where $E \setminus H \subseteq V \setminus H$ which is a null set. We also have $E = V \setminus (V \setminus E)$ where $V \setminus E \subseteq V \setminus H$, which is null. ∎

Recall: If E is measurable, then we define its *Lebesgue measure* by $m(E) = m_*(E)$.

Theorem 2: The Lebesgue measure is countably additive, i.e.

$$E_i \bigcap E_j = \emptyset \implies m(\bigcup E_i) = \sum m(E_i).$$

Proof: Assume each E_j is bounded, so that $m_*(E_j) < \infty$. Given $\varepsilon > 0$, for each j we can find a *compact* K_j such that

$$m(E_j \setminus K_j) \leq \varepsilon/2^j.$$

Then for any finite N, since the E_j are disjoint, then $\{K_i\}_{i=1}^N$ are also disjoint, so

$$m(\bigcup^N K_j) = \sum m(K_j).$$

However, we have $m(E_k) - m(K_j) < \varepsilon/2^j$, and so $m(K_j) > m(E_j) - \varepsilon/2^j$. Then

$$m(\bigcup^N K_j) = \sum m(K_j) \geq \sum m(E_j) - \varepsilon/2^j = \sum m(E_j) - \varepsilon.$$

But since

$$\bigcup^N K_j \subset E := \bigcup^\infty E_j,$$

we have

$$\begin{aligned}
m(E) \geq m(\bigcup^N K_j) \geq \sum^N m(E_j) \implies \\
\sum^n m(E_j) \leq m(E) + \varepsilon \forall N \implies \\
\sum^\infty m(E_j) \leq m(E) + \varepsilon && \to m(E).
\end{aligned}$$

So this shows the bounded case.

In general, let

$$\begin{aligned}
A_1 &= B(1,0) \\
A_2 &= B(2,0) \setminus B(1,0) \\
&\vdots
\end{aligned}$$

Then let $E_{i,j} = E_i \bigcap A_j$, so $\bigcup_i E_i = \bigcup_{i,j} E_{i,j}$, where all of the $E_{i,j}$ are still disjoint but also now bounded.

Then

$$m(\bigcup E_j) = m(\coprod_{i,j} E_{i,j}) \qquad = \sum_j \sum_i m(E_{i,j}) = \sum_j m(E_i),$$

where the last two equalities follow from the bounded case. ■

11 Tuesday September 10th:

11.1 A Brief Introduction to (Actual) Measure Theory

Definition: Instead of just $\mathbb{R}^n$, just consider a set X with a σ-algebra $\mathcal{A}$.

Then the pair $(X, \mathcal{A})$ is a *measurable* space, i.e. it is ready to be equipped with a measure.

A *measure* space is a measurable space with a *measure*.

Definition: A set function $\mu : \mathcal{A} \to [0, \infty]$ satisfying

- $\mu(\emptyset) = 0$
- μ is countably additive,

is said to be a *measure*.

> What we've done so far is construct something we've called "the Lebesgue measure", then verified that it was actually a measure. We constructed a set function on $\mathbb{R}^n$ called m (the outer measure), then restricted attention to a class of sets $\mathcal{M}$, and produced a measure space $(\mathbb{R}^n, \mathcal{M}, m)$.

Note that countable additivity implies monotonicity and subadditivity, which were what we originally discovered about m_*.

11.2 Continuity of Measure

Theorem (Fundamental Theorem of Measure Theory):

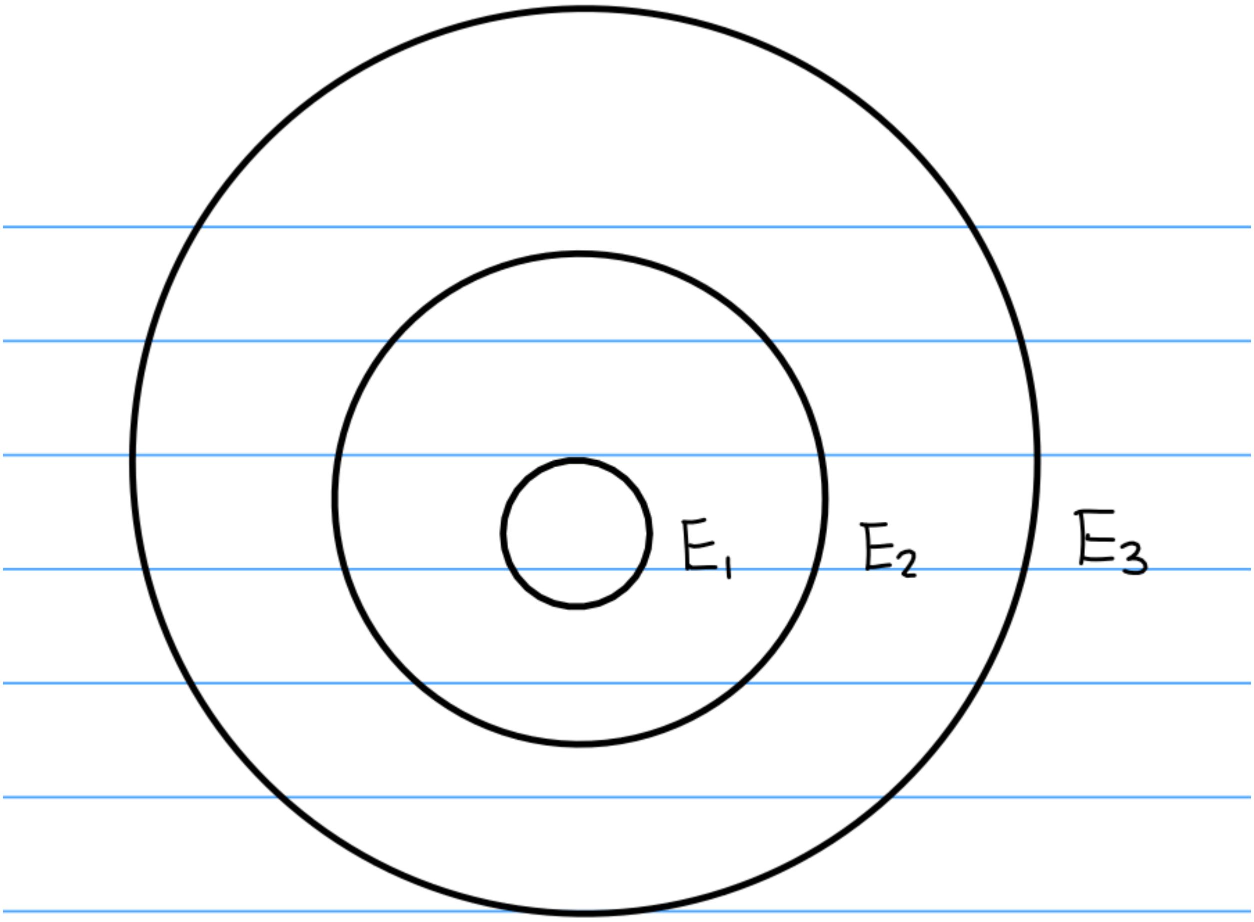
Let $\{E_i\} \subseteq \mathcal{M}$.

- If $E_i \nearrow E$, so $E_1 \subseteq E_2 \subseteq \cdots$ and $\bigcup E_i = E$, then $\mu(E) = \lim \mu(E_i)$ (continuity from below).
- If $E_i \searrow E$, so $E_1 \supseteq E_2 \supseteq \cdots$ and $\bigcap E_i = E$, then $\mu(E) = \lim \mu(E_i)$ as long as $\mu(E_1) < \infty$i (continuity from above).

Exercise: Show that $\mu(E_1) < \infty$ is necessary in the 2nd result above.

Proof:

Part (1):



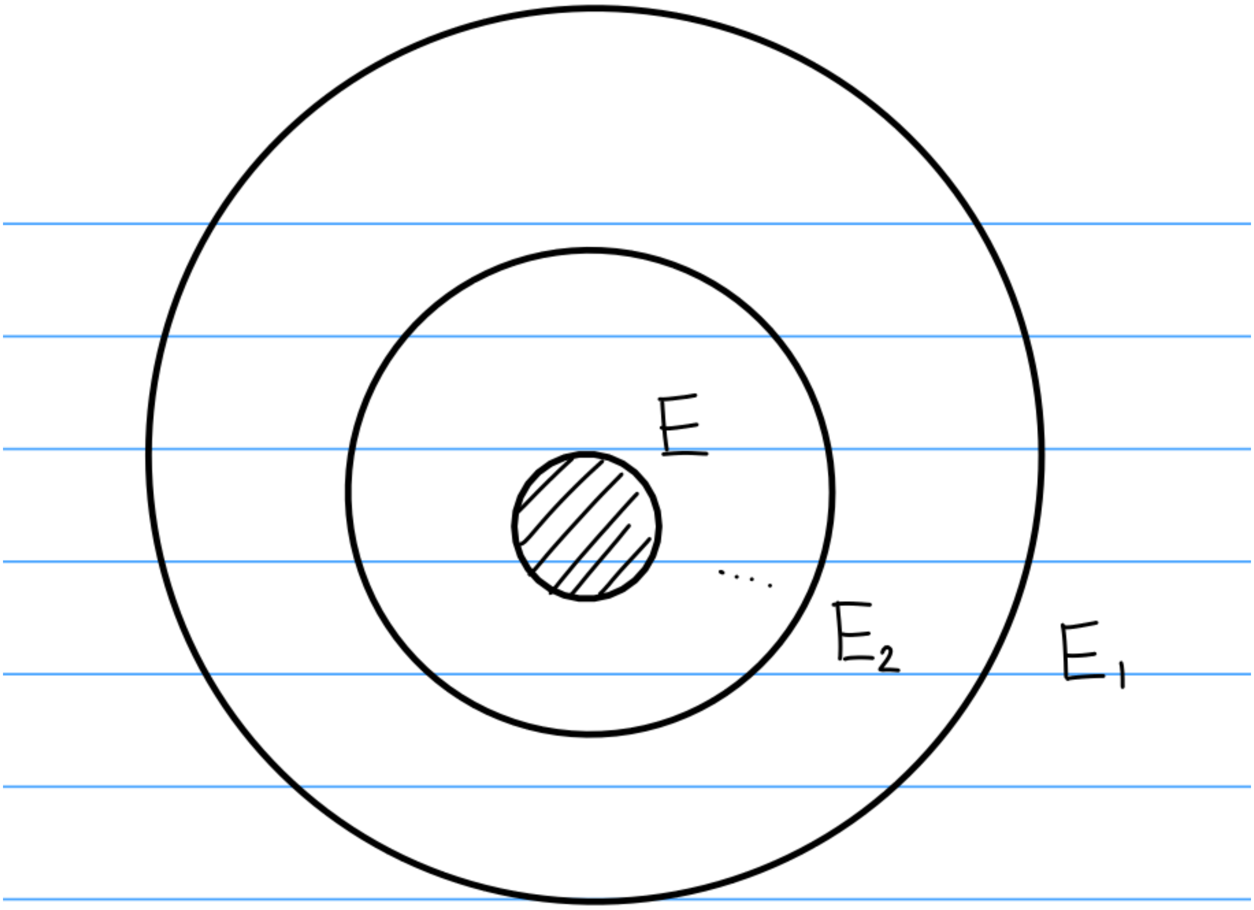
Let

- $A_1 = E_1$
- $A_2 = E_2 \setminus E_1$,
- $\cdots A_j = E_j \setminus E_{j-1}$.

Then $\{A_j\}$ are disjoint, and $E = \coprod A_j$, so

$$\begin{aligned}
m(E) &= \sum_j m(A_j) \\
&= \lim_{k\to\infty} \sum_{j=1}^{k} \mu(A_j) \\
&= \lim_{k\to\infty} \mu(\bigcup_{j=1}^{k} A_j) \\
&= \lim_{k\to\infty} \mu(E_k). \quad \blacksquare
\end{aligned}$$

Part (2):



Let $A_i = E_j \setminus E_{j+1}$, so $\{A_i\}$ are disjoint.

Then (**important!!**) $E_1 = \bigcup A_i \bigcup E$, which are **disjoint and measurable**. Then,

$$
\begin{aligned}
\mu(E_1) &= \sum \mu(A_i) + \mu(E) \\
&= \lim_{k\to\infty} \sum_{i=1}^{k-1} \mu(A_j) \\
&= \lim_{k\to\infty} \sum_{i=1}^{k-1} \mu(E_i) - \mu(E_{i+1}) + \mu(E), \quad \text{(which is telescoping)} \\
&= \lim_{k\to\infty} (\mu(E_1) - \mu(E_k)) + \mu(E) \\
\implies \mu(E_1) &= \mu(E_1) - \lim_k \mu(E_k) + \mu(E) \quad \text{since } \mu(E_1) < \infty \\
\implies \mu(E) &= \lim_k \mu(E_k).
\end{aligned}
$$

■

Recall that if $E \subseteq \mathbb{R}^n$, then

$$
m_*(E) = \inf \left\{ m_*(G) \;\middle|\; E \subset G \text{ open} \right\}
$$
$$
\iff
$$
$$
\forall \varepsilon > 0,\ \exists G \supseteq E \;\middle|\; m_*(G) \le m_*(E) + \varepsilon.
$$

Note: this says that the measure is *regular*.

11.3 Approximation by Compact Sets

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

$$
m(E) = \sup \left\{ m(K) \;\middle|\; K \subseteq E \text{ and } K \text{ is compact} \right\}
$$
$$
\iff
$$
$$
\forall \varepsilon\ \exists \text{ compact } K \subseteq E \;\middle|\; m(K) \ge m(E) - \varepsilon.
$$

Proof:

Case 1: Suppose E is bounded.

Let $\varepsilon > 0$, then by the closed characterization of measurability, we have $m(E \setminus F) < \varepsilon$ for some closed set $F \subseteq E$.

Since E is bounded, $m(E) < \infty$, and so

$$
\begin{aligned}
m(E \setminus F) < \varepsilon &\implies \\
m(E) - m(F) < \varepsilon &\implies \\
m(F) &> m(E) - \varepsilon.
\end{aligned}
$$

where since F is a closed and bounded set in $\mathbb{R}^n$, F is compact.

Case 2: Suppose E is unbounded. Write

$$E_j = E \bigcap_{j=1}^{\infty} B(j, 0),$$

so $E_j \nearrow E$.

Using continuity from below, we have $m(E) = \lim_j m(E_j)$.

Suppose $m(E) < \infty$. Then for some j,

$$m(E_j) \geq m(E) - \varepsilon.$$

But E_j is bounded.

By case 1, there is a compact $K \subseteq E_j$ where

$$m(K) \geq m(E_j) - \varepsilon.$$

But then $m(K) \geq m(E) - 2\varepsilon$.

Suppose now that $m(E) = \infty$.

For any $M > 0$, we can find an E_j such that

$$m(E_j) > M.$$

Since E_j is bounded, by case 1 we get a compact $K \subseteq E_j \subseteq E$ with $m(K) > M$.

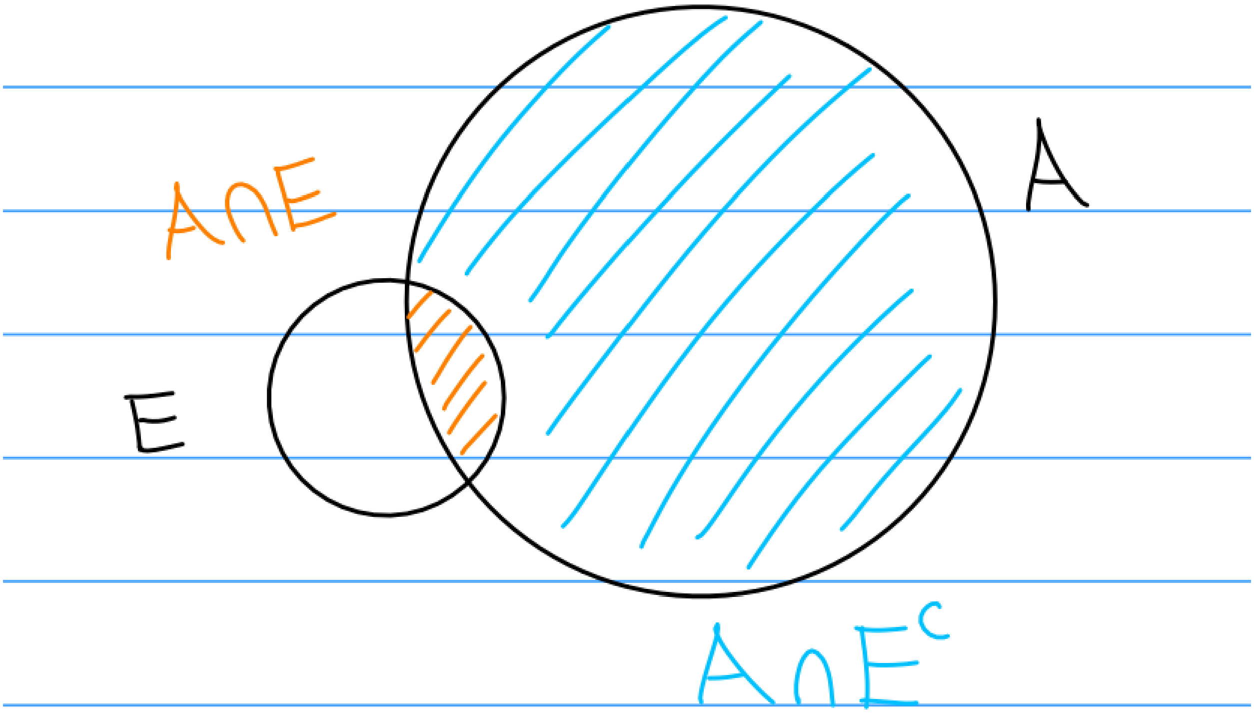
■

Qual alert: very similar arguments are often used on the quals.

11.4 Caratheodory Characterization

A subset $E \subseteq \mathbb{R}^n$ is measurable iff for all $A \subseteq \mathbb{R}^n$,

$$m_*(A) = m_*(E \bigcap A) + m_*(E \bigcap A^c).$$



Note that this can be interpreted in terms of *inner measures*, in which we're approximating E with cubes from the inside. If we also think of this in terms of probability spaces, we could interpret the RHS as saying that the probability of an event either happening or *not* happening should be 1.

Note: G_δ sets are measurable.

Note that the $\leq$ case is satisfied automatically by subadditivity, and the $\geq$ case comes from if $A \subseteq V$ then

$$m(A) \geq \cdots \text{Exercise!}.$$

Theorem: Let $E \subseteq \mathbb{R}^n$ be measurable. Then

1. For all $h \in \mathbb{R}^n$, then $E + h$ is measurable, and $\mu(E + h) = \mu(E)$.
2. For every $x \in \mathbb{R}$, the set cE is measurable and $\mu(cE) = |c|^n \mu(E)$.

We can say more, and determine measures of all linear transformations of a set in terms of the determinant. Note that because we're working with cubes in the outer measure, so the only content here is that these new sets are actually measurable.

If E is measurable, $E = H \bigcup Z$ where $H \in F_\sigma$ and $\mu(Z) = 0$. But then $E + h = (H + h) \bigcup (Z + h)$, but $H + h$ is still F_σ because shifts of closed sets are still closed. Moreover, $\mu(Z + h) = \mu(Z) = 0$, so were done.

Moral: it suffices to check things on Borel sets.

12 Thursday September 12th

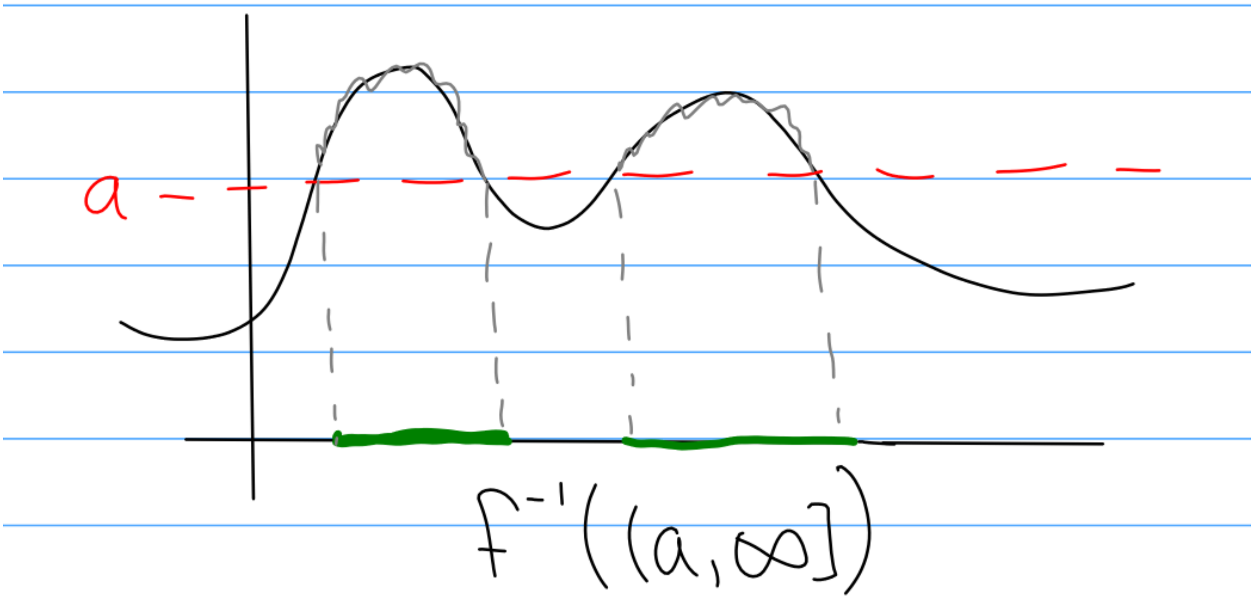
12.1 Measurable Functions

Let $E \subseteq \mathbb{R}^n$ be measurable. Then $f : E \to \mathbb{R} \bigcup \{\pm\infty\}$ is *Lebesgue measurable* iff

$$\left\{x \in E \;\middle|\; f(x) > a\right\} = f^{-1}((a, \infty])) \in \mathcal{M},$$

the collection of Lebesgue measurable sets, for every $a \in \mathbb{R}$.

Similarly, E is Borel measurable if we replace $\mathcal{M}$ by $\mathcal{B}$, the collection of all Borel measurable sets.



As usual, there are many different characterizations:

- $f^{-1}((a, \infty])) \in \mathcal{M} \quad \forall a \in \mathbb{R}$, since we can write this as $\bigcap f^{-1}((a - \dfrac{1}{k}, \infty])$, which are all measurable.
- $f^{-1}([a, \infty]) \in \mathcal{M} \quad \forall a \in \mathbb{R}$, by taking complements
- $f^{-1}([-\infty, a]) \in \mathcal{M} \quad \forall a \in \mathbb{R}$.

Theorem: If $f : E \to \mathbb{R}$ is finite-valued, then

$$f \text{ is measurable} \iff f^{-1}(G) \in \mathcal{M} \text{ for all open } G \subseteq \mathbb{R}.$$

Proof:

$\Longleftarrow$: Easy, since (a, ∞) is always an open set.

$\Longrightarrow$: Suppose f is measurable and let $G \subseteq \mathbb{R}$ be open, then $G = \coprod I_i$ where each I_i is an interval. Then $f^{-1}(G) = \bigcup f^{-1}(I_i)$.

But if $I_i = (a,b)$, then $f^{-1}(I_i) = f^{-1}((a,\infty)) \bigcap f^{-1}((-\infty,b))$, both of which are measurable by definition.

12.2 Extending the Class of Measurable Functions

Corollary: Continuous functions are in fact Borel measurable, and in particular Lebesgue measurable.

Proposition: If $f : E \to \mathbb{R}$ is measurable and $\varphi : \mathbb{R} \to \mathbb{R}$ is continuous, then $\varphi \circ f$ is measurable.

> Note: This condition can not be relaxed to just φ being measurable. However, this does work if you require that φ is Borel measurable.

Proof: Let $G \subseteq \mathbb{R}$ be open. Then

$$(\varphi \circ f)^{-1}(G) = f^{-1}(\varphi^{-1}(G)) \in \mathcal{M}.$$

But $\varphi^{-1}(G)$ is open since φ is continuous, and thus measurable, and since f is a measurable function, this pulls back to a measurable set, so the composition is measurable.

Consequences of this proposition: If f is measurable, then so is $|f|, |f|^p, f^2, e^{cf}$, for any constant, etc.

We will show that $\mathcal{M}$ is closed under most algebraic and limiting operations.

Theorem 2: If f, g are $\mathbb{R}$-valued measurable functions, then fg and $f+g$ are measurable.

> Note that we already know this for g a constant, since $x \mapsto x+c$ and $x \mapsto cx$ are continuous functions.

Proof: To come later.

An interesting consequence: if f, g are measurable then $\max(f,g)$ is as well, since

$$\max(f,g) = \frac{(f+g) + |f-g|}{2}.$$

Theorem: If $\{f_n\}$ is a sequence of $\overline{\mathbb{R}}$-valued measurable functions, then

1. $\sup_n f_n$ and $\inf_n f_n$ are measurable
2. $\limsup_n f_n \coloneqq \inf_n \sup_{k \geq n} f_k$ is measurable, as is the $\liminf_n f_n$.

> Note: As a consequence, if $f_n \to f$ pointwise and each f_n is measurable, then $\limsup f_n = \lim f_n = f$ is measurable.

Proof of Theorem 2:

Suppose f, g are measurable and finite-valued. Let $a \in \mathbb{R}$ and consider

$$S = \left\{ x \;\middle|\; f(x) + g(x) > a \right\}.$$

Then

$$S = \left\{ x \;\middle|\; f(x) > -g(x) + a \right\},$$

where f is measurable and so is $-g+a$.

With the following lemma, we'll be done:

Lemma: If f, h are measurable, then $S = \left\{x \mid f(x) > h(x)\right\}$ is always a measurable set.

Proof: Since $f(x) > h(x)$, there is a rational q such that $h(x) < q < f(x)$.

But then

$$S = \bigcup_{q \in \mathbb{Q}} \left\{x \mid f(x) > q > h(x)\right\} = \{f > q\} \bigcap \{h < q\},$$

which is an intersection of measurable sets. ■

For the set equality, just check that $x \in S \implies x \in \bigcup \text{stuff}$, and $x \notin S \implies x \notin \bigcup \text{stuff}$.

Note: we can write $fg = \dfrac{(f+g)^2 - (f-g)^2}{4}$.

Proof of Theorem 3:

Since $\inf_n f_n(x) = -\sup f_n(x)$, we only need to show that $\sup f_n$ is measurable.

For any given a, we want to show that $S = \left\{x \mid \sup_n f_n(x) > a\right\}$ is measurable.

Then there is some n for which $f_n(x) > a$, and the claim is that

$$S = \bigcup_n \left\{x \mid f_n(x) > a\right\}.$$

But this follows formally by just checking set inclusions. So S is a countable union of measurable sets and thus measurable. ■

12.3 Almost Everywhere Equality

Definition: If $f, g : E \to \overline{\mathbb{R}}, \mathbb{C}$ then $f = g$ *almost everywhere* (or $f = g$ a.e.) iff

$$\left\{x \mid f(x) \neq g(x)\right\} \text{ is null }.$$

Fact: If f is measurable and $f = g$ a.e., then g is measurable.

This follows because

$$\{g > a\} = \{f > g\} \bigcup Z,$$

where $Z \subseteq \{f \neq g\}$, which is null, forcing Z to be null as well.

Fact: If $\{f_n\}$ is a sequence of measurable functions and $f_n \to f$ pointwise a.e., then f is measurable.

Note that we've replaced open, continuous, uniform etc with a new notion of "measurable" in all instances.

This new notion isn't so far from the original ones, but allows much more to be done.

Littlewood's Principles:

- Every measurable set is *nearly open* (See *Lebesgue density*)
- Every measurable function is *nearly continuous* (See *Lusin's Theorem*)
- Every convergent sequence of measurable functions is *nearly uniformly convergent* (See *Egorov's Theorem*)

13 Tuesday September 17th?

13.1 Approximation by Simple Functions

Definition: Let $E \subseteq \mathbb{R}^n$ be measurable. Then the *characteristic function* of E is defined as

$$\chi_E(x) \coloneqq \begin{cases} 1 & x \in E \\ 0 & \text{else.} \end{cases}$$

Definition: A *step function* is a function of the form

$$S(x) = \sum_{i=1}^{N} a_i \chi_R(x)$$

where R is some rectangle.

Definition: A *simple function* is a function of the form

$$s(x) = \sum_{i=1}^{N} a_j \chi_{E_j}(x)$$

where each E_j is measurable.

Theorem 1: If $f : E \to [0, \infty]$ is a non-negative measurable function, then there exists a sequence of simple functions $\{s_k\}$ such that

$$s_k(x) \leq s_{k+1}(x) \ \forall x, k \quad \text{and} \quad \lim_{k \to \infty} s_k(x) = f(x) \ \forall x.$$

Corollary: This in fact holds for $f : E \to \overline{\mathbb{R}}$ taking on extended real values, not just positive functions.

Proof: Write $f = f^+ - f^-$, where $f^+(x) = \max\{f(x), 0\}$.

Theorem 2: If $f : E \to \overline{\mathbb{R}}$ is measurable, there exists a sequence ψ_k of *step* functions such that $\psi_k \to f$ almost everywhere.

Proof: See homework 3, problem 1c.

13.2 Lebesgue Density Theorem

Theorem (Lebesgue Density): If $E \subseteq \mathbb{R}^n$ is measurable, then

$$\lim_{r \to 0^+} \frac{m(E \bigcap B(r, x))}{m(B(r, x))} = 1 \quad \text{for almost every } x \in E.$$

13.3 Egorov's Theorem

Theorem (Egorov): Let $E \subseteq \mathbb{R}^n$ be measurable with $m(E) > 0$.

Let $f_k : E \to \mathbb{R}$ be a sequence of measurable functions such that $f(x) := \lim_{k\to\infty} f_k(x)$ exists a.e. and is finite valued.

Then the convergence is *almost uniform*, i.e.

$$\forall \varepsilon > 0,\ \exists F \subseteq E \text{ closed } \Bigm| m(E \setminus F) < \varepsilon \text{ and } f_k \rightrightarrows f \text{ on } F.$$

Are these conditions really necessary?

1. If $E = \mathbb{R}$, let
$$f_k(x) = \mathbb{1}\left[|x| > k\right].$$
Then $f_k \to 0$ a.e. but not "almost uniformly".
2. If $E = [0, 1]$, let
$$f_k(x) = k\mathbb{1}\left[0 \leq x \leq 1 - \frac{1}{k}\right].$$
Then
$$f_k \to \infty \iff 0 \leq x < 1,\ 0 \iff x = 1,$$
but not "almost uniformly".

13.4 Lusin's Theorem

Theorem (Lusin):

Suppose f is measurable and finite-valued on a measurable set E with $m(E) < \infty$.

Then $\forall \varepsilon > 0,\ \exists F \subseteq E$ closed such that

$$m(E \setminus F) < \varepsilon \text{ and } f|_F \text{ is continuous }.$$

> This doesn't mean that the original f is actually continuous on F, when thought of as a function on E – we restrict the universe to only F, so e.g. we can only take sequences that are subsets of F when we go to check continuity. Example to note:
>
> $$f = \chi_{\mathbb{Q}} \bigcap [0, 1],$$
>
> which is discontinuous at every point.

13.5 Proof of Egorov

Proof of Egorov: Assume wlog that $f_k \to f$ *everywhere.*

Lemma: Let $E, \{f_n\}$ and f be as in Egorov's theorem.

Then for all $\varepsilon > 0, \alpha > 0$ there exists a closed set $F \subseteq E$ and some $k_0 \in \mathbb{N}$ such that

- $m(E \setminus F) < \varepsilon$, and
- $|f_k(x) - f(x)| < \alpha \quad \forall x \in F,\ k \geq k_0$.

So let $\varepsilon > 0$, then the lemma tells us that for every j we can find a closed set $F_j \subseteq E$ with

$$m(E \setminus F_j) < \varepsilon/2^j$$

and k_j such that

$$|f_k(x) - f(x)| < \frac{1}{j} \text{ on } F_j \quad \forall k \geq k_j.$$

So take $F := \bigcap F_j$, which is *closed.*

By subadditivity, we have

$$m(E/F) \leq \sum m(E \setminus F_j) < \varepsilon$$

by construction.

Note that the convergence is *uniform*, since k_j in the lemma already provided the uniform threshold for all points in F_j, and $F \subseteq F_j$ for every j. ■

Proof of lemma: Fix ε, α.

Define

$$\begin{aligned} E_j &:= \left\{ x \in E \;\middle|\; |f_k(x) - f(x)| < \alpha \ \forall k > j \right\} \\ &= \bigcap_{k=j+1}^{\infty} \left\{ x \in E \;\middle|\; |f_k(x) - f(x)| < \alpha \right\}. \end{aligned}$$

Note $E_j \subseteq E_{j+1}$ and $E = \bigcup E_j$, so we have $E_j \nearrow E$.

Using continuity from below, $\lim_j m(E_j) = m(E)$.

Since $m(E) < \infty$, there exists a k_0 such that

$$j \geq k_0 \implies m(E \setminus E_j) < \varepsilon/2.$$

So choose $F \subset E_{k_0}$ be closed with the property that

$$m(E_{k_0} \setminus F) \leq \frac{\varepsilon}{2}.$$

So if $x \in F$, then $x \in E_{k_0}$ and thus $x \in E_j$ for all $j \geq k_0$ since they are nested.

But then

$$k \leq k_0 \implies |f_k(x) - f(x)| < \alpha,$$

and we're done. ■

13.6 Convergence in Measure

Definition: Let $E \subseteq \mathbb{R}^n$ be measurable and $f, \{f_k\}$ be measurable, finite-valued functions defined on E.

We say that $f_k \to^m f$ or $f_k \to f$ *in measure* if for every $\alpha > 0$, we have

$$\lim_{k \to \infty} m(\left\{ x \in E \;\middle|\; |f_k(x) - f(x)| > \alpha \right\}) = 0.$$

■

How does this relate to pointwise convergence?

Theorem: If $m(E) < \infty$, then

$$f_k \to f \text{ } a.e \text{ on } E \iff f_k \to^m f \text{ on } E.$$

Proof: Exercise using the previous lemma.

> Note that the converse is false! See homework exercise. There is a partial converse: convergence in measure will yield a *subsequence* that converges almost everywhere.

14 Thursday September 19th

14.1 Review of the Lebesgue Integral for L^+

Recall the definition of L^+, and the fact that for any $f \in L^+$, there is a sequence $\{f_n\}$ of simple functions in L^+ such that $f_n \nearrow f$.

Given any simple function ϕ, we defined

$$\int \phi = \sum_j a_j m(E_j)$$

iff this is a standard representation for ϕ.

We then extend to all functions in L^+ by defining

$$\int f := \sup \left\{ \int \phi \;\middle|\; 0 \leq \phi \leq f, \; \phi \text{ simple} \right\}.$$

Some properties:

- $f \leq g \implies \int f \leq \int g$

  (monotonicity, easy to show)

- $\displaystyle\int cf = c\int f$

  (also easy to show)

- $\displaystyle\int (f+g) = \int f + \int g$

  (less obvious, follows from MCT)

Theorem: If $\{f_n\} \subset L^+$, then $\sum \int f_k = \int \sum f_k$.

Proof: Exercise, not too tricky.

14.2 Proof of Monotone Convergence Theorem

Theorem (MCT): If $\{f_n\} \subset L^+$ with $f_k \leq f_{k+1}$ and $f_k \to f$, then $\lim \int f_n = \int \lim f_n = \int f$.

Proof of MCT: Given any simple $\phi \in L^+$, define the set function

$$\begin{aligned} \mu_\phi : \mathcal{M} &\to [0, \infty] \\ A &\mapsto \int_A \phi. \end{aligned}$$

So if $\{E_k\} \subset \mathcal{M}$ and $E_1 \subseteq E_2 \subseteq \cdots$, then $\mu_\phi(\bigcup E_k) = \lim \mu_\phi(E_k)$.

> Note that
>
> $$f_k \in L^+ \implies \lim f_k \in L^+ \quad \text{and} \quad f_k \leq f_{k+1} \implies \int f_k \leq \int f_{k+1},$$
>
> so the limit on the LHS makes sense.

Let $f = \lim f_k$, then

$$\int f_k \leq \int f \text{ for all } k,$$

so

$$\lim \int f_k \leq \int f.$$

So we need to show that $\lim \int f_k \geq \int f$.

Fix $\alpha \in (0,1)$ and let ϕ be simple with $0 \leq \phi \leq f$. (We'll later show that the result is independent of this choice.)

Let

$$E_k = \left\{ x \;\middle|\; f_k(x) \geq \alpha\phi(x) \right\},$$

which we can now say is *clearly* measurable.

Then $E_1 \subseteq E_2 \subseteq \cdots$ since the f_k are increasing. Moreover, $\bigcup E_k = \mathbb{R}^n$ (check!), since $f_k \to f, \alpha\phi < \phi \leq f$.

Then

$$\int f_g \geq \int_{E_k} \geq \alpha \int_{E_k} \phi := \alpha\mu_\phi(E_k).$$

But then

$$\lim \int f_k \geq \lim \alpha\mu_\phi(E_k) = \alpha\mu_\phi(\mathbb{R}^n) = \alpha \int \phi \quad \forall\alpha, \forall\phi.$$

So $\lim \int f_k \geq \int \phi$ for all ϕ simple with $0 \leq \phi \leq f$.

Thus

$$\int f = \sup_\phi \int \phi \leq \lim \int f_k,$$

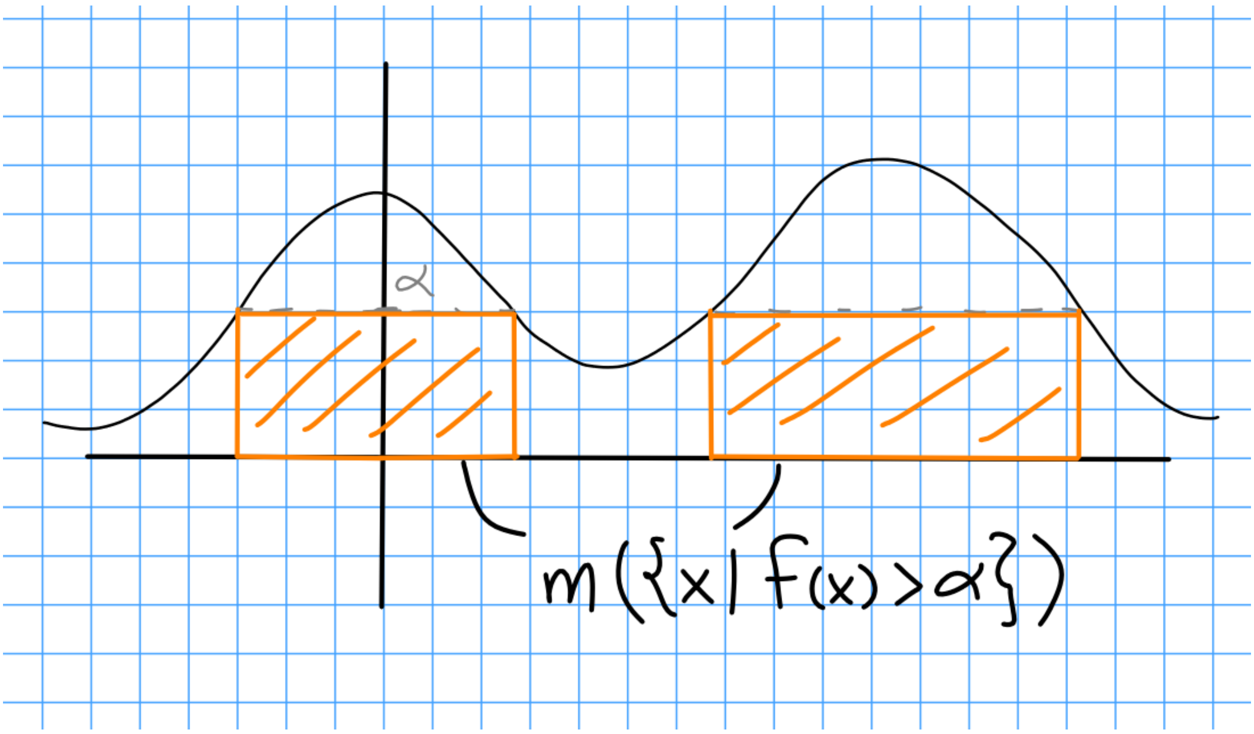
which is what we wanted to show. ■

14.2.1 Chebyshev's Inequality

Theorem (Chebyshev's Inequality): If $f \in L^+$, then

$$m(\left\{x \mid f(x) \geq \alpha\right\}) \geq \frac{1}{\alpha} \int f \forall\alpha.$$

Proof:



Just note that

$$\alpha m(\left\{x \;\middle|\; f(x) > \alpha\right\}) = \int \alpha \chi_{f \geq \alpha} \leq \int f.$$

■

14.2.2 The Integral Detects Almost-Everywhere Equality

Proposition: Suppose $f \in L^{+}$. Then

$$\int f = 0 \implies f = 0 \text{ a.e.}$$

> Nice trick: showing $a_n \to L$ can be done by showing $a_n - L \to 0$. Similarly,
>
> $$\int f = \int g \iff \int (f - g) = 0$$

Proof of Proposition:

This is obvious for simple functions:

If $f = \sum a_i \chi_{E_i}$ in standard representation, then

$$\int f = \sum a_j m(E_j) = 0 \iff \text{ either } a_j = 0 \;\forall j, \text{ or } a_j \neq 0 \text{ and } m(E_j) = 0.$$

So it only disagrees with zero on a measure zero set.

In general, suppose $f = 0$ a.e., note that

$$\phi \leq f = 0 \implies \phi = 0 \text{ a.e.}$$

and thus $\int \phi = 0$ by the previous case.

But then

$$\int f = \sup_{\phi} \int \phi = 0 \text{ a.e. },$$

and we're done.

Now suppose $\int f = 0$ a.e.; we can apply Chebyshev,which says that

$$m(\left\{x \;\middle|\; f(x) \geq \alpha\right\}) = 0 \text{ for any } \alpha \geq 0.$$

But then

$$\begin{aligned}
m(\left\{x \;\middle|\; f(x) > 0\right\}) &= m(\bigcup_n \left\{x \;\middle|\; f(x) > \frac{1}{n}\right\}) \\
&= \lim m(\text{stuff}) \\
&= \lim 0 \\
&= 0.
\end{aligned}$$

■

Remark: In the MCT, the monotonicity is necessary, i.e. we really need $f_k \nearrow f$.

Examples:

- $f_k = k\chi_{[0,\frac{1}{k}]}$. Then $f_k \to 0$ a.e. but $\int f_k = 1$ for all k while $\int f = 0$.
- $f_k = \chi_{[k,k+1]} \to 0$ a.e. (the skateboard to infinity!)

14.3 Fatou's Theorem

Another convergence theorem, this time with virtually no hypothesis:

Theorem (Fatou): If $\{f_k\} \subset L^+$, then $\int \liminf f_k \leq \liminf \int f_k$

How to remember: in the above examples, we had $\int \lim f_k = 0$, so we can saturate the LHS to zero to obtain an inequality of the form $a \leq b$.

Proof of Fatou: We can write

$$\liminf_k f_k = \lim_k g_k \quad \text{where} \quad g_k = \inf_{n \geq k} f_n.$$

Note that $g_k \nearrow \liminf_k f_k$, we can apply MCT. So

$$\begin{aligned} \int \liminf_k f_k &= \int \lim_k g_k \\ &=_{\text{MCT}} \lim_k \int g_k \\ &= \liminf \int g_k \\ &\leq \liminf \int f_k, \end{aligned}$$

where we can note that

$$g_k \leq f_k \implies \int g_k \leq \int f_k.$$

■

14.4 Dominated Convergence Theorem

Theorem (DCT): If $\{f_k\} \subset L^+$, $f_k \to f$ a.e., **and** $f_k \leq g \in L^+$ where $\int g < \infty$, then

$$\int f = \int \lim f_k = \lim \int f_k.$$

15 Tuesday September 24th

15.1 Convergence Theorems

Two main convergence theorems: define $L^+ = \left\{ f : \mathbb{R}^n \to [0, \infty] \;\middle|\; f \text{ is measurable.} \right\}$. Then

Theorem 1 (MCT): If $\{f_n\} \subseteq L^+$ with $f_n \nearrow f$, then

$$\int f = \lim \int f_n.$$

Theorem 2 (Fatou's lemma): If $\{f_n\} \subseteq L^+$, then

$$\int \liminf f_n \leq \liminf \int f_n.$$

Corollary 1: If $\{f_n\} \subseteq L^+$ and $f_n \to f$ a.e. and $\displaystyle\int f_n \leq M$ uniformly, then

$$\int f \leq M$$

This uses Fatou's lemma.

Corollary 2: If $\{f_n\} \subseteq L^+$ and $f_n \to f$ a.e. with instead $f_n \leq f$ a.e. for all n, then

$$\int f = \lim \int f_n.$$

Proof: By Fatou,

$$\int f \leq \liminf \int f_n.$$

If we can show $\displaystyle\limsup \int f_n \leq \int f$ as well, we're done.

Since integrals satisfy monotonicity,

$$f_n \leq f \implies \int f_n \leq \int f \text{ a.e.}$$

But by order-limit laws, we then have $\displaystyle\limsup \int f_n \leq \int f$ as well. ■

15.2 Extending the Integral to $\overline{\mathbb{R}}$-valued functions (and $\mathbb{C}$)

Definition: A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is *integrable* iff $\displaystyle\int |f| < \infty$. Note that $|f| = f_+ + f_-$, so if $f : \overline{\mathbb{R}} \to \mathbb{R}$ is integrable, then

$$\int |f| = \int (f_+ + f_-) = \int f_+ + \int f_-,$$

which means both must be finite. We now have two finite *numbers*, so we can subtract.

Definition: For $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, we define

$$\int f = \int f_+ - \int f_-.$$

Similarly, for $f : \mathbb{C} \to \overline{\mathbb{R}}$, let $f = \mathrm{Im} f + i \mathrm{Re} f$, and define

$$\int f = \int \Re f + i \int \Im f.$$

Note: the space of all $\overline{\mathbb{R}}$ or $\mathbb{C}$ valued functions forms a real (resp. complex) vector space, and the integral is a real (complex) linear functional. This is not immediate – multiplying by scalars is clear, but distributing integrals over sums is not. If $h = f + g$, then it is **not** the case that $h_+ = f_+ + g_+$. But we can write out

$$h = f + g \implies h_+ + h_- = f_+ + f_- + g_+ + g_-,$$

maneuver things so that everything is positive, and *then* take integrals.

15.3 L^1 and its Norm

Definition: We can provisionally define

$$L^1 = \{ f : \mathbb{R}^n \to \mathbb{C} : f \text{ is measurable } \},$$

where we'd like to define

$$\|f\|_{L^1} \coloneqq \int |f|.$$

However, this only yields a *seminorm*, since nonzero functions still end up with zero norm. This can be remedied by identifying functions which agree on a set of measure zero.

Proposition (Triangle Inequality for L^1 Seminorm): If $f \in L^1$, then $\left| \int f \right| \leq \int |f|$.

Proof: This is trivial if $\int f = 0$. If f is $\mathbb{R}$-valued, then

$$\begin{aligned}
\left| \int f \right| &= \left| \int f_+ - \int f_- \right| \\
&\leq \left| \int f_+ \right| + \left| \int f_- \right| \\
&= \int f_+ + \int f_- \\
&= \int f_+ + f_- \\
&= \int |f|.
\end{aligned}$$

If f is $\mathbb{C}$-valued, then $|z| = \dfrac{z^* z}{|z|}$, so

$$\begin{aligned}
\frac{(\int f)^*}{|\int f|} \int f &:= \alpha \int f \\
&= \int \alpha f \\
&= \int \operatorname{Re}(\alpha f) + i \int \operatorname{Im}(\alpha f).
\end{aligned}$$

but since what we started with was *real*, the imaginary component vanishes, so this equals

$$\int \operatorname{Re}(\alpha f) = \left| \int \operatorname{Re}(\alpha f) \right| \leq \int |\operatorname{Re}(\alpha f)| \leq \int |\alpha f| = \int |f|.$$

> Note: this is referred to as a **rotation/triangle trick**.

Actual definition of L^1: Let $X \subseteq \mathbb{R}^n$ be measurable. Then

$$L^1(X) = \{\text{equivalence classes of a.e. defined integrable functions on } X\}$$

This is an equivalence relation, and we write $f \sim g \iff f = g$ a.e.

We then define $\|f\|_1 = \displaystyle\int_X |f|$.

We have

$$\int |f| = 0 \text{ a.e.} \iff f = 0 \text{ a.e.}$$

Exercise: Prove this for L^+ functions.

We'd like this to also be true iff $\displaystyle\int_E f = 0$ for all measurable $E \subseteq X$.

15.4 Dominated Convergence Theorem

The following theorem will apply whenever we want to switch integrals and limits, but f is *not necessarily* in L^+.

> This is the ONLY theorem that doesn't require non-negativity!!!

Theorem (DCT): Suppose $\{f_n\} \in L^1$, $f_n \to f$ a.e., and $|f_n| \leq g$ a.e. for all n with $g \in L^1$.

Then $f \in L^1$ and

$$\int f = \lim \int f_n \quad \text{and} \quad \lim \int |f_n - f| = 0$$

> Note that the second statement is stronger, and in fact implies the first. This statement is what we'll prove.

Proof: Since $f_n \to f$ and each f_n is measurable, then f is measurable. Since $f_n \leq g$ for all n, then $f \leq g$. So

$$\int |f| \leq \int |g| < \infty,$$

and thus $f \in L^1$.

It suffices to show that $\limsup \int |f_n - f| \leq 0$.

We have to get something non-negative to apply anything we know so far, so

$$\begin{aligned}
0 \leq |f_n - f| &\leq |f_n| + |f| \leq g + |f| \\
&\implies g + |f| - |f_n - f| \geq 0 \\
&\text{(This is where we'll apply Fatou.)} \\
&\implies \int \liminf(g + |f| - |f_n - f|) \leq \liminf \int (g + |f| - |f_n - f|) \\
&\implies \int (g + |f|) \leq \liminf \int (g + |f|) - \liminf \int |f_n - f| \\
&\implies \int (g + |f|) \leq \int (g + |f|) + \limsup \int |f_n - f| \\
&\implies 0 \leq \limsup \int |f_n - f|. \blacksquare
\end{aligned}$$

$\blacksquare$

15.5 Differentiating Under the Integral

Let

$$F(t) = \int f(x,t)\ dx$$

- Is F continuous at a point t_0?
- Is F differentiable at t_0?

We could show continuity by looking at

$$\lim_{t \to t_0} |F(t) - F(t_0)| \leq \lim_{t \to t_0} \int |f(x,t) - f(x,t_0)| \leq_{DCT} \int \lim |f(x,t) - f(x,t_0)|.$$

which will go to zero exactly when f is continuous in t.

Differentiability can be shown by considering

$$\begin{aligned}
\lim \frac{|F(t_0) - F(t)|}{t - t_0} &\leq \lim \int \frac{f(x,t) - f(x,t_0)}{t - t_0}\ dx \\
&\leq \int \lim \frac{f(x,t) - f(x,t_0)}{t - t_0}\ dx \\
&= \int f'(x,t_0)\ dx.
\end{aligned}$$

16 Thursday September 26th

16.1 L^1 and its Convergence Theorems

For any measurable $X \subseteq \mathbb{R}^n$, we defined

$$L^1(X) = \left\{ f : X \to \mathbb{C} \text{ measurable } \;\middle|\; \int_X |f| < \infty \right\} / \sim$$

where $f \sim g \iff f = g$ a.e.

> Note that we could talk about $\overline{\mathbb{R}}$ valued functions, *but* (theorem) integrable functions can only be finite on a null set. So we can stop considering these altogether if we're just considering L^1 functions.

The space L^1 is in fact a *normed vector space* with

$$\|f\|_{L^1(X)} := \int_X |f|.$$

> Recall that we needed to identify functions because this was only a *seminorm* otherwise, and we only want the zero function to have norm zero.

We say

$$f_n \xrightarrow{L^1} f \iff \|f_n - f\|_1 \to 0.$$

Convergence Theorems:

> Mantra: Everything positive and some positivity: MCT. More often: DCT.

- **MCT**:

$$f_n \in L^+, \quad f_n \nearrow f \text{ a.e. } \implies \lim \int f_n = \int f.$$

  - Note that it's very important that $f_n \in L^+$
  - **Corollary**:

$$\sum \int f_n = \int \sum f_n.$$

- **DCT:**

$$f_n \in L^1, \quad f_n \to f \text{ a.e. }, \quad |f|_n \leq g \in L^1 \implies \lim \int f_n = \int f.$$

  - **A Stronger statement**:

$$f_n \xrightarrow{L^1} f \quad \text{i.e.} \int |f_n - f| \to 0.$$

  The previous statement only gives $\left| \int f_n - f \right| \to 0$. This follows because

$$\lim \int |f_n - f| =_{DCT} \int \lim |f_n - f| \to 0,$$

  since $|f_n - f| \leq 2g$.

16.2 Commuting Sums with Integrals

Theorem: If

- $f_n \in L^1$, and
- $\sum_n \int |f|_n < \infty$,

Then $\sum_n f_n$ converges to an L^1 function and

$$\sum_n \int f_n = \int \sum f_n.$$

Note that uniform convergences $\implies$ pointwise $\implies$ a.e. convergence, and so we should think of convergence in norm as *weaker* than all of these (although they are not actually comparable).

Proof: By the MCT, we know

$$\int \sum |f|_n =_{MCT}= \sum \int |f|_n,$$

which is integrable, and so the first term is integrable as well.

By the homework problem,

$$\sum |f|_n \in L^1 \implies \sum |f_n(x)| < \infty \text{ for almost every } x.$$

So consider just these x values.

Note that "$\mathbb{R}$ is complete" is equivalent to "absolutely convergent implies convergent" for sums.

So for each x, $\sum f_n(x)$ converges. What are the partial sums?

$$\left|\sum^N f_j(x)\right| \leq \sum^\infty |f_j(x)| \quad \forall j, \ a.e. \ x.$$

So let $g_N = \sum^N f_j$, so g_N is dominated by $g := g_\infty$. Then

$$\begin{aligned}
\int \sum_j^\infty f_j &= \int \lim_{n\to\infty} \sum_{j=1}^n f_j \\
&=_{DCT} \lim_{n\to\infty} \sum_{j=1}^n \int f_j \\
&= \sum_{j=1}^\infty \int f_j.
\end{aligned}$$

■

Note that these partial sums are converging a.e., **and in** L^1. We didn't use this here, but it will be important when we want to show that L^1 is complete.

16.3 Different Notions of Convergence

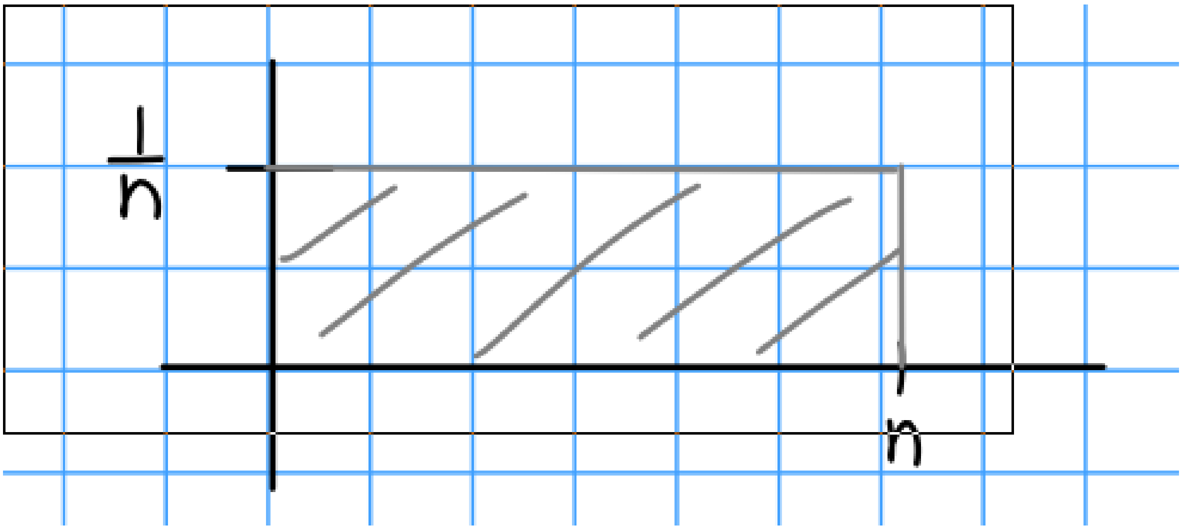
Note that $f_n \to f$ can mean many things:

1. Uniform: $f_n \rightrightarrows f : \forall \varepsilon\ \exists N \ \Big|\ n \geq N \implies |f_N(x) - f(x)| < \varepsilon \quad \forall x.$
2. Pointwise: $f_n(x) \to f(x)$ for all x. (This is just a sequence of numbers)
3. Almost Everywhere: $f_n(x) \to f(x)$ for almost all x.
4. Norm: $\|f_n - f\|_1 = \displaystyle\int |f_n(x) - f(x)| \to 0.$

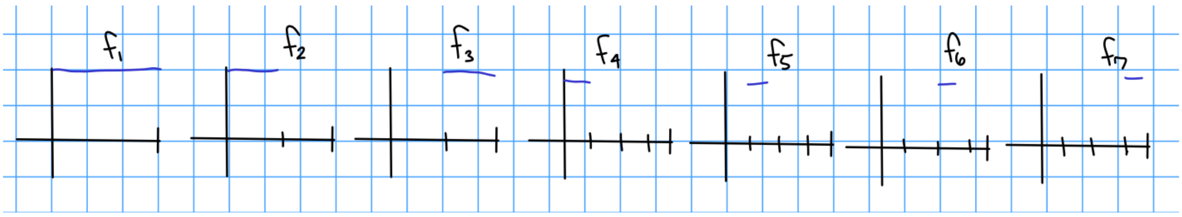
We have $1 \implies 2 \implies 3$, and in general no implication can be reversed, but (**warning**) none of $1, 2, 3$ imply 4 or vice versa.

Examples:

- $f_n = n^{-1}\chi_{[0,n]}$. This converges uniformly to 0, but the integral is identically 1. So this satisfies 1,2,3 and not 4.



- $f_n = \chi_{(n,n+1)}$ (skateboard to infinity). This satisfies 2,3 but not 1, 4.
- $f_n = n\chi_{(0,\frac{1}{n})}$. This satisfies 3 but not 1,2,4.
- f_n : see weird example below. Then $f_n \to 0$ in L^1 but is not 1,2, or 3.



16.4 Comparing L^1 Convergence to a.e. Convergence

Theorem: If $f_n \to f \in L^1$, then there is a subsequence f_{n_k} such that $f_{n_k} \to f$ almost everywhere.

Note: convergence always implies Cauchy, so we'll assume this right away.

Since f_n converges in L^1, it is Cauchy in L^1, so $\|f_n - f_m\|_1 \to 0$.

Note: we want to pick a sequence that is converging *faster* when we construct our subsequence, since that's the obstruction to a.e. convergence.

So there is a subsequence $n_1, n_2, \cdots$ such that $\|f_n - f_m\| \leq 2^{-k}$ if $n, m \geq n_k$. Let $g_1 = f_{n_2}, g_k = f_{n_{k+1}} - f_{n_k}$ be the consecutive differences. Then

- $\|g_{k+1}\| \leq 2^{-k}$ for all k,
- $f_{n_k} = \sum_{j=1}^{k} g_j$

Thus we want to show that this sum converges almost everywhere to an L^1 function. So if $\sum_{j=1}^{\infty} \|g_j\|_1 < \infty$, we're done.

We have

$$\sum \|g_j\| = \|g_1\| + \sum_j 2^{-j}.$$

By the previous theorem, this means $f_{n_k} = \sum^{k} g_j \xrightarrow{a.e.} f$.

We know it converges to *some* L^1 function, **but limits are unique**, so this is actually the original f. ■

16.5 Completeness of L^1

Theorem: L^1 is a complete normed space, i.e. **a Banach space**, so every Cauchy sequence in L^1 converges to a function in L^1.

Proof:

Proofs of completeness tend to go the same way:

1. Take a Cauchy sequence $\{f_n\}$.
2. Find a candidate limit f
3. Show that the f_n actually converge to this candidate f
4. Show that f is in L^1.

So suppose f_n is Cauchy.

From the previous theorem, we know a subsequence (all in L^1) converges to some limit f in L^1. So let this f be the candidate limit, we just need to show that $\|f_n - f\|_1 \to 0$.

Let $\varepsilon > 0$ and choose k large enough such that

- $2^{-k} \leq \frac{1}{2}\varepsilon$.
- $\|f_{n_k} - f\|_1 \leq \varepsilon$.

Then

$$\begin{aligned}\|f_n - f\|_1 &\le \|f_n - f_{n_k}\|_1 + \|f_{n_k} - f\|_1 \\ &\le 2^{-k} + \varepsilon/2 \\ &\le \varepsilon/2 + \varepsilon/2 = \varepsilon.\end{aligned}$$

■

17 Tuesday October 1

17.1 Completeness of L^1 (Revisited)

Last time: L^1 is complete, where we used the fact that $\mathbb{R}$ is complete in the following way

Theorem:

$$\mathbb{R} \text{ is complete} \iff \left(\sum |x|_n \infty \implies \sum x_n < \infty\right).$$

Proof:

$\implies$:
Suppose $\mathbb{R}$ is complete and $\sum |x|_n < \infty$.

Let $S_N = \sum_{i=1}^{N} x_n$. Then if $N > M$,

$$|S_N - S_M| \le \sum_{i=M+1}^{N} |x|_n \to 0.$$

$\impliedby$: Suppose every absolutely convergent series is convergent.

Let $\{x_n\}$ be Cauchy; we want to show that it is convergent as well.

> Note: we'll use the same trick as last time. The goal is to cook up an absolutely convergent sequence, the convergence of which will imply convergence of our original series.

Choose a subsequence

$$n_1 \le n_2 \le \cdots \quad \text{such that} \quad |x_n - x_m| < 2^{-j} \quad \text{if} \quad n, m \ge n_j.$$

Let $y_1 = x_{n_1}$ and $y_j = x_{n_j} - x_{n_{j-1}}$ for $j > 1$.

Then

$$x_{n_k} = \sum_{i=1}^{k} y_j \quad \text{and} \quad \sum_{j=1}^{\infty} |y_j| \le |y|_1 + \sum_{j=2}^{\infty} 2^{-k} < \infty.$$

So $\lim x_{n_k}$ exists and equals $\sum y_j$.

It follows that for $n > n_k$ and k is sufficiently large,

$$|x_n - x| \leq |x_n - x_{n_k}| + |x_{n_k} - x| < \varepsilon.$$

■

Theorem (Modified): Let X be a normed vector space.

$$X \text{ is complete } \iff \left(\sum_n \|x_n\| < \infty \implies \sum_n x_n < \infty \right).$$

Proof: Completely the same, just replace absolute values with norms everywhere!

17.2 Translation and Dilation Invariance of the Lebesgue Integral

Qual Problem Alert!

Definition: Define a *translation* $\tau_h(x) := x + h$ and $\tau f(x) := f(x - h)$ for all $h \in \mathbb{R}^\times$.

Definition: Define a *dilation* $f_\delta(x) := \delta^{-n} f(\delta^{-1} x)$ for all $\delta > 0$.

Theorem:

1.

$$f \in L^1 \implies \tau_h f \in L^1 \text{ and } \int \tau_h f = \int f$$
$$\left(\text{i.e. } \int_E f(x - h) = \int_{E+h} f \right).$$

2.

$$f \in L^1 \implies f_\delta \in L^1 \text{ and } \int f_\delta = \int f$$
$$\left(\text{i.e. } \delta^{-n} \int f(\delta^{-1} x) = \int f(y) \right).$$

Proof: We first verify this for $f = \chi_E$ where $E \in \mathcal{M}$.

We have $\tau_h f(x) = f(x - h) = \chi_E(x - h) = \chi_{E+h}(x)$ and

$$\int \tau_n f = m(E + h) = m(E) = \int f,$$

where we know the measures are equal by translation invariance of measure.

By linearity, this holds for simple functions as well.

Useful technique: once you know something for simple functions, you can often apply MCT to get it for L^+ functions as well!

If now $f \in L^+$ then there exists a sequence of simple functions $\{\phi_k\} \nearrow f$, and by the MCT,

$$\int \phi_k \to \int f.$$

Note that

$$\{\tau_h \phi_k\} \nearrow \tau_h f,$$

so

$$\int \tau_h \phi_k \to \int \tau_h f.$$

But $\left\{\int \tau_h \phi_k\right\} = \left\{\int \phi_k\right\}$, so **by uniqueness of limits** we must have

$$\int f = \int \tau_h f.$$

Now this follow for $\mathbb{R}$-valued functions by writing $f = f_+ - f_-$, and then for $\mathbb{C}$-valued functions by $f = \Re(f) + i\Im(f)$.

$\blacksquare$

17.3 Agreement of Riemann and Lebesgue Integrals

Theorem: Let f be a bounded $\mathbb{R}$-valued function on a closed interval $[a, b]$.

If f is Riemann integrable, then $f \in L^1$ (so $\mathcal{R} \subseteq L^1$ is a subspace) and the integrals agree, so

$$\int_a^b f(x)\ dx = \int_{[a,b]} f(x)\ dx$$

Proof: Given a partition $P = \{t_1, t_2, \cdots t_n\}$ of $[a, b]$, let

$$\begin{aligned} G_p &= \sum_{j=1}^n \sup \left\{ f(x) \;\middle|\; x \in [t_k, t_{j+1}] \right\} \chi_{[t_j, t_{j+1}]} \\ g_p &= \sum_{j=1}^n \inf \left\{ f(x) \;\middle|\; x \in [t_k, t_{j+1}] \right\} \chi_{[t_j, t_{j+1}]}. \end{aligned}$$

Then $\int G_p = U(f, P)$ and $\int g_p = L(f, p)$ where U, L denote the upper and lower sums. Note that the Riemann integral is the infimum of the former and the supremum of the latter, over increasingly fine partitions.

So let $\{P_k\}$ be a sequence of partitions with the size of the mesh going to 0.

Then $G_{P_k} \searrow G$ is converging to something, and $g_{p_k} \nearrow g$. In particular, we have

$$G_P \leq f \leq g_P \quad \text{and so} \quad G \leq f \leq g.$$

Since f is bounded, say by M, then both of these sequences are dominated by $\pm M\chi_{[a,b]} \in L^1$. So we can invoke the DCT, which yields

$$\int G_{P_k} \to \int G \text{ and } \int g_{P_k} \to \int g,$$

and thus

$$\int g = \int G = \int_{\mathcal{R}} f.$$

Since

$$\int G = \int g \implies \int (G-g) = 0 \implies G = g \text{ a.e. },$$

we have $f = G$ a.e.

But G is a sequence of measurable function, and so f is measurable. Moreover, $\displaystyle\int G = \int f$. But $\displaystyle\int G = \int_{\mathcal{R}} f$ as well, so the two integrals agree.

> Recall that
>
> $$f = 0 \text{ a.e.} \iff \int_E f = 0 \text{ for all } E \subseteq \mathcal{M}.$$

■

Examples next time:

Continuous functions with compact support are dense in L^1, which is a version of the following:

Littlewood's Principle: Any integrable function is *almost* continuous, in the sense that for any f and any $\varepsilon > 0$ there is a continuous function g such that

$$\int f - \int g < \varepsilon.$$

18 Thursday October 3

18.1 Relating Zero Functions to Zero Integrals Over Measurable Sets

Theorem: $f = 0$ a.e. iff $\displaystyle\int_E f = 0$ for all $E \in \mathcal{M}$.

If $f \in L^+$ we already know that $f = 0$ a.e. iff $\displaystyle\int f = 0$.

$\implies$: Since $f = 0$ a.e., we have $|f| = 0$ a.e. and since $|f| \in L^+$ we have $\displaystyle\int |f| = 0$.

Now let $E \in \mathcal{M}$; then

$$\left|\int_E f\right| \leq \int_E |f| \leq \int |f| = 0.$$

$\Longleftarrow$: Suppose $\displaystyle\int_E f = 0 \; \forall E \in \mathcal{M}$ and $f \neq 0$ a.e., then either

1. $f+$ is positive on a set of nonzero measure, or
2. f^- is positive on a set of nonzero measure.

So suppose wlog (1) holds.

Let $E = \left\{ x \;\middle|\; f^+ > 0 \right\}$, then $m(E) > 0$.

Then $\displaystyle\int_E f^+ > 0$, since $\chi_E f^+ \neq 0$ almost everywhere.

We also know that $f^+ \in L^+$, so

$$f^+ = 0 \text{ a.e.} \quad \Longleftrightarrow \quad \int f = 0.$$

But then $\displaystyle\int_E f > 0$, since

$$\text{support}(f^+) \bigcap \text{support}(f^-) = \left\{ x \;\middle|\; f(x) = 0 \right\},$$

so $f^- = 0$ on E.

■

18.2 Approximation Theorems and Dense Subspaces of L^1

Definition: We say that a collection $\mathcal{C}$ of functions is *dense* in L^1 iff

$$\forall \varepsilon > 0 \quad \text{and} \quad \forall f \in L^1, \quad \exists g \in \mathcal{C} \quad \text{such that} \quad \|f - g\|_1 < \varepsilon.$$

Theorem(s):

1. Simple functions are dense in L^1. **(DCT)**
2. Continuous functions with compact support (C_c or C_0) are dense in L^1.
3. Step functions are dense in L^1.

Proof of (1): Let $f \in L^1$ and $\varepsilon > 0$.

Since f is measurable, there exists a sequence of simple functions $\{\phi_k\} \to f$ pointwise with $|\phi_k| \leq |\phi_{k+1}|$. Then f dominates ϕ_k and the DCT yields $\displaystyle\int |\phi_k - f| < \varepsilon$ for k large enough.

■

> We'll use this as a stepping stone – we really want to get *continuous functions*, but now we can show there are continuous functions arbitrarily close to *simple* functions, and the triangle inequality will give us the desired result.

Proof of (2): We have shown that there exists a simple function $\phi = \displaystyle\sum_{j=1}^N a_j \chi_{E_j}$ in standard representation, where $a_j \neq 0$ and the E_j are disjoint, with $\displaystyle\int |f - \phi| < \varepsilon$.

It suffices to show that for all j, there exists a $g_j \in C_c$ such that $\left\|\chi_{E_j} - g_j\right\| < \varepsilon$.

Note that if we have this, we can define $g = \sum a_j g_j \in C_c$.

But then

$$\int |\phi - g| = \int \left|\sum_i^N a_i(\chi_{E_j} - g_j)\right| \leq \sum_i^N |a_i| \left|\chi_{E_j} - g_j\right| \leq C\varepsilon.$$

Then applying the triangle inequality yields the desired result.

Important Observation:

Each E_j has finite measure, so we have

$$m(E_j) = \frac{1}{|a_j|} \int_{E_j} |\phi| \leq \frac{1}{|a_j|} \int |\phi| < \infty.$$

Claim: If $m(E) < \infty$, then there exists a $g \in C_c$ such that $\|\chi_E - g\|_1 < \varepsilon$ for any $\varepsilon > 0$.

Proof: Note that we can find a $K \subseteq E \subseteq G$ such that K is compact, G is open, and $m(G \setminus K) < \varepsilon$.

Since K is closed and G^c is closed, **by Urysohn's Lemma**, there is a continuous g such that $\chi_K \leq g \leq \chi_G$. But then g is zero on G^c and 1 on K.

Then $|\chi_E - g|$ is supported on $G \setminus K$, so

$$\int |\chi_E - g| \leq m(G \setminus K) < \varepsilon.$$

■

Remark: We will eventually show that *smooth* compactly supported functions are also dense in L^1.

This approximation theorem yields some nice proofs:

18.3 Small Tails and Absolute Continuity

Proposition: If $f \in L^1$ and $\varepsilon > 0$, then

1. **Small tails**:

$$\exists N \text{ such that } \int_{\|x\| \geq N} |f| < \varepsilon$$

   Take $f_N = f\chi_{B(N)} \nearrow f$

2. **Absolute continuity:** There exists a $\delta > 0$ such that

$$m(E) < \delta \implies \int_E |F| < \varepsilon$$

   Take $f_N = f\chi_S$ *where* $S = \{f(x) \leq N\}$.

Useful technique: If you want to prove something for L^1 functions, try to show it's true for C_c functions.

Note that we know $\exists g \in C_c$ such that $\int |f-g| < \varepsilon$.

Proof of (1): Let N be large enough such that $g = 0$ if $|x| \geq N$. Let $E = \left\{x \;\middle|\; |x| \geq N\right\}$.

Then

$$\int_E |f| = \int_E |f-g+g| \leq \int_E |f-g| + \int_E |g| < \varepsilon + 0.$$

Proof of (2): There exists an M such that $|g| \leq M$, since C_c functions are bounded almost everywhere.

Then

$$\int_E |g| \leq M \cdot m(E) < \varepsilon.$$

So set $\delta = \varepsilon/M$, then if $m(E) < \delta$ then

$$\int_E |f| \leq \int |f-g| + \int_E |g| < \varepsilon.$$

18.4 Continuity in L^1

Qual problem alert: Prove the following theorem. Note that DCT doesn't quite work!

Theorem (Continuity in L^1)

$$f \in L^1 \implies \lim_{h\to 0} \int |f(x+h) - f(x)| = 0.$$

Proof: Let $\varepsilon > 0$.

Then choose g such that

$$\int |f(x) - g(x)| < \varepsilon.$$

By translation invariance, $\int |f(x+h) - g(x+h)| < \varepsilon$ as well.

Qual problem alert: remember how to prove translation invariance of the Lebesgue integral.

Now

$$\int |f(x+h) - f(x)| \leq 2\varepsilon + \int |g(x+h) - g(x)|.$$

Since g is continuous and has compact support, g is uniformly continuous.

So enlarge the support of g to a compact set K such that $|g(x+h)-g(x)|=0$ for all $x \in K^c$ and $|h| \leq 1$. But then

$$\int_K |g(x+h)-g(x)| \leq \varepsilon \int_K 1 \to 0.$$

Note that $\operatorname{supp}(F) = \overline{\left\{x \;\middle|\; f(x) \neq 0\right\}}$.

19 Tuesday, October 8

Notation: think of $\mathbb{R}^n = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ where $n_1 + n_2 = n$.

Motivation: If $f(x,y)$ is measurable, is it true that $f(y) := f(x_0, y)$ also measurable for a fixed x_0?

19.1 Fubini and Tonelli

Theorem (Tonelli): Let $f(x,y)$ be non-negative and measurable on $\mathbb{R}^n$. Then for almost every $x \in \mathbb{R}^{n_1}$, we have

1. $$f_x(y) := f(x,y)$$

   is measurable as a function of y in $\mathbb{R}^{n_2}$

2. $$F(x) := \int f(x,y)\ dy$$

   is measurable as a function of x,

3. $$G(y) = \int F(x)\ dx = \int \left(\int f(x,y)\ dy \right)\ dx$$

   is measurable and equal to $\displaystyle\int_{\mathbb{R}^n} f$.

Corollary: If $E \subset \mathcal{M}\left(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}\right)$, then for a.e. $x \in \mathbb{R}^{n_1}$, the slice

$$E_x := \left\{y \in \mathbb{R}^{n_2} \;\middle|\; (x,y) \in E\right\}$$

is measurable.

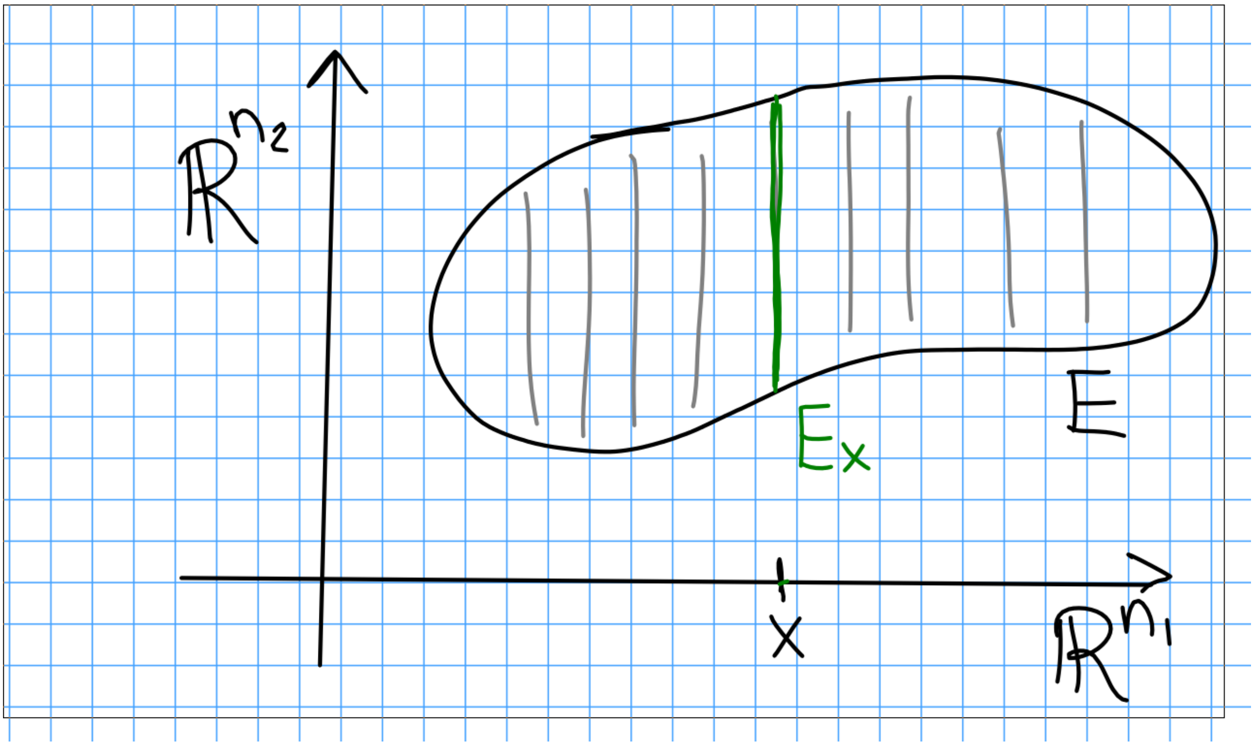
Moreover, $x \mapsto m(E_x)$ is a measurable function of x, and

$$m(E) = \int_{\mathbb{R}^n} m(E_x)\ dx.$$

Warning: We assumed E was measurable here, but it is possible for every slice to be measurable while E itself is not!

Take $E = \mathcal{N} \times I$ for $\mathcal{N}$ the unmeasurable set. Then $E_x = \chi_{[0,1]}$ and so the image is always measurable.

But taking y slices yields $E_y \chi_{\mathcal{N}}$, which (by the above corollary) would have to be measurable if E were measurable.



Note: We need to show that taking a **cylinder** on a function (i.e. given $f(x)$ and defining $F(x, y) = f(x)$) does not destroy measurability. This is necessary in the context of convolution, since $f(x - y)$ will need to be measurable in both variables in order to apply Tonelli.

19.2 Application: Area Under the Graph

Suppose $f \geq 0$ on $\mathbb{R}^n$, with no assumption of measurability.

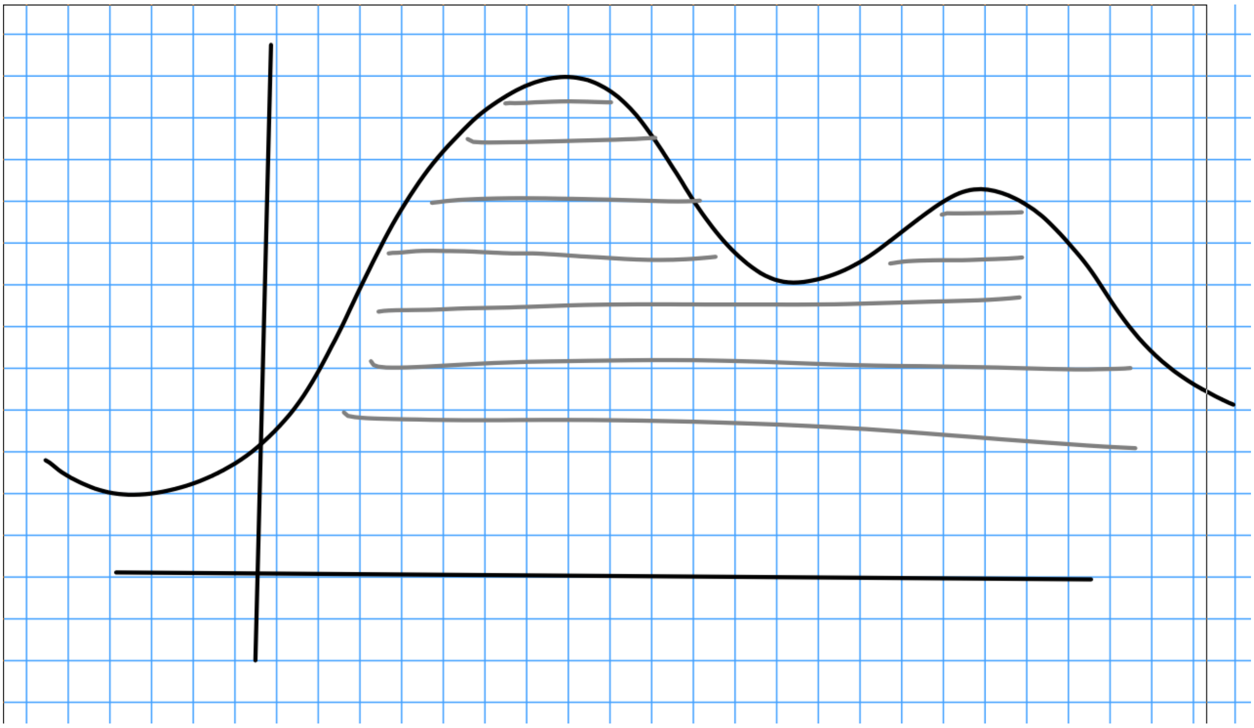
Consider defining the "area under the graph" as

$$\mathcal{A} := \left\{ (x, y) \in \mathbb{R}^n \times \mathbb{R} \;\middle|\; 0 \leq y \leq f(x) \right\}.$$

Then

1. f is measurable on $\mathbb{R}^n$ iff $\mathcal{A}$ is a measurable subset of $\mathbb{R}^{n+1}$.
2. If f is measurable on $\mathbb{R}^n$, then

$$m(\mathcal{A}) = \int_{\mathbb{R}^n} f = \int_0^\infty m(\left\{ x \in \mathbb{R}^n \;\middle|\; f(x) \geq y \right\}) \, dy$$



Proof of (1):

$\implies$: Suppose f is measurable on $\mathbb{R}^n$.

By the lemma, $F(x, y) = f(x)$ is measurable on $\mathbb{R}^n \times \mathbb{R}$ and $G(x, y) = y$ is as well. But then $\mathcal{A} = \{G \leq F\} \bigcap \{G \geq 0\}$, which is an intersection of measurable sets and thus measurable.

$\impliedby$: Suppose $\mathcal{A}$ is measurable.

By Tonelli, for almost every $x \in \mathbb{R}^n$, the slice

$$\mathcal{A}_x = \left\{ y \in \mathbb{R} \;\middle|\; (x, y) \in \mathcal{A} \right\} = [0, f(x)]$$

is measurable.

Then $m(\mathcal{A}_x) = f(x)$, so $x \mapsto \mathcal{A}_x$ is a measurable function of x and $m(\mathcal{A}) = \displaystyle\int f(x)\, dx$.

Repeating this argument with y slices instead, for almost every $y \in \mathbb{R}$ we have

$$\mathcal{A}_y = \left\{ x \in \mathbb{R}^n \;\middle|\; (x, y) \in \mathcal{A} \right\} = \left\{ x \in \mathbb{R}^n \;\middle|\; f(x) \geq y \geq 0 \right\},$$

which is a measurable subset of $\mathbb{R}^n$.

So it makes sense to integrate it, and

$$m(\mathcal{A}) = \int m(\mathcal{A}_y)\, dy = \int_0^y m(\left\{ x \in \mathbb{R}^n \;\middle|\; f(x) \geq y \right\})\, dy.$$

$\blacksquare$

Alternative proof:

$$\int_0^\infty m(\left\{x \in \mathbb{R}^n \;\middle|\; f(x) \geq y\right\}) = \int_{\mathbb{R}} \int_{\mathbb{R}^n} \chi_{S := \left\{x \in \mathbb{R}^n \;\middle|\; f(x) \geq y \geq 0\right\}}$$
$$= \int_{\mathbb{R}^n} \int_{\mathbb{R}} \chi_S$$
$$= \int_{\mathbb{R}^n} \int_0^{f(x)} \, dy \, dx$$
$$= \int_{\mathbb{R}^n} f(x) \, dx.$$

■

19.3 Appendix on Measurability in $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$.

Lemma: If f is measurable on $\mathbb{R}^{n_1}$, then $F(x, y) := f(x)$ is measurable on the product space $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$

Qual problem alert.

Proof of Lemma: Suppose f is measurable on $\mathbb{R}^n$; we want to show that $F(x, y) = f(x)$ is measurable on $\mathbb{R}^n \times \mathbb{R}$.

This amounts to showing that for any a,

$$S_a := \left\{(x, y) \;\middle|\; F(x, y) \geq a\right\} \in \mathcal{M}(\mathbb{R}^{n+1}).$$

But we can rewrite

$$S_a = \left\{x \in \mathbb{R}^n \;\middle|\; f(x) > a\right\} \times \mathbb{R},$$

which is the cylinder on a measurable set. As we will show, this is always measurable.

Best way to show measurability: use Borel characterization, or show that it's an $H \coprod N$ where $H \in F_\sigma$ and N is null.

So write $E = H \coprod N$ where $H \in F_\sigma$ and N is null.

Then $E \times \mathbb{R} = (H \times \mathbb{R}) \bigcup (N \times \mathbb{R})$.

But $H \times \mathbb{R}$ is still an F_σ set, so we just need to show $N \times \mathbb{R}$ is still null.

We have $N \times [-k, k] \nearrow N \times \mathbb{R}$, so we can use continuity from below.

To see that $m(N \times [-k, k]) = 0$, first cover N by such that $\sum |Q_i| < \varepsilon/2k$.

But the measure of any rectangle over such a cube will be $M(\overline{Q}_i) = 2k \cdot m(Q_i)$, which we can pull out of $\sum \left|\overline{Q}_i\right|$. ■

19.4 Fubini and Fubini-Tonelli

Summary":

- **Tonelli**: *Non-negative and measurable* allows switching integrals,
- **Fubini**: *Just measurable* allows switching the integrals, the integrals are finite, and all iterated variants are equal.
- **Fubini/Tonelli**: Extends switching beyond just non-negative integrands.

Theorem (Fubini): Let $f(x, y)$ be an integrable function on $\mathbb{R}^n \times \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$.

Then for almost every $x \in \mathbb{R}^{n_1}$,

1. $f_x(y) = f(x, y)$ is an *integrable* function of y in $\mathbb{R}^{n_2}$.
2. $\displaystyle\int_{\mathbb{R}^{n_2}} f(x, y)\ dy$ is an integrable function of x in $\mathbb{R}^{n_1}$.

Moreover,

$$\int_{\mathbb{R}^n} f = \int \int f(x, y)\ dx\ dy$$

in either order.

Theorem (Fubini-Tonelli): Let $f(x, y)$ be measurable in the product space.

If either

$$\int \left(\int |f(x, y)\ dy| \right)\ dx < \infty$$

or

$$\int \left(\int |f(x, y)\ dx| \right)\ dy < \infty$$

then by Tonelli on $|f(x, y)|$, we can conclude $f \in L^1(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$.

Moreover, by Fubini, $\displaystyle\int f$ is equal to either iterated integral.

> Moral: If *any* iterated integral is finite, then they *all* are.

Comparing this to sums: recall that $\displaystyle\sum \int f_n = \int \sum f_n$ is true exactly when

1. $f_n \geq 0$, and
2. $\displaystyle\sum \int |f_n| < \infty$.

20 Tuesday October 15

20.1 Proof of Fubini's Theorem

Recall the strong version of DCT: It allows you to deduce L^1 converges from a.e. convergence.

Note that otherwise, this are incomparable!

Proof of Fubini's Theorem: see book for the gory details.

> Essentially uses MCT a number of times and reduces to the case of cubes that possibly include boundaries.

21 Thursday October 17

21.1 Review of Tonelli

Theorem (Tonelli): Suppose $f(x,y)$ is non-negative on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ and measurable on the product space. Then for a.e. x, we have

1. $f_x(y)$ is a measurable function of y.
2. Since it's non-negative and measurable, the integral makes sense, and

$$\int_{\mathbb{R}^{n_2}} f(x,y)\ dy$$

is a measurable function of x.

3.

$$\int\int f(x,y)\ dx\ dy = \int\int f(x,y)\ dy\ dx = \int f.$$

Proof:

> Qual problem alert: useful technique!

Essentially use Fubini, and truncate domain/range with a k-ball.

See proof of the *case-jumping lemma* and notes on webpage for details.

> We'll never need to dig into the proof of this, but there will *always* be a question related to *applying* it.

21.2 Measurability of Linear Transformations

Theorem: Let $T \in \mathrm{GL}(n, \mathbb{R})$. Then

- f measurable $\implies$ $f \circ T$ is measurable.
  - Contrast to what happens for g a continuous function instead of T.

$$f \leq 0 \text{ or } f \in L^1 \implies \int f = |\det(T)| \int (f \circ T)(x).$$

-

$$E \in \mathcal{M}(\mathbb{R}^n) \implies T(E) \in \mathcal{M}(\mathbb{R}^n).$$

It suffices to prove this for Borel sets and Borel measurable functions. This can be proved using Fubini.

Note that if we choose f to be Borel measurable, then $f \circ T$ will be measurable because T is continuous.

> This follows because $\left\{ E \;\middle|\; T^{-1}(E) \in \mathcal{B} \right\}$ is in fact a σ-algebra that contains all open sets. *Exercise*: Prove this.

We can also reduce this to proving the result for T_i an elementary matrix, since if it holds for T, S then it holds for TS because

$$\int f = \det T \int f \circ T = \det T \det S \int f \circ T \circ S = \det(TS) \int f \circ (TS)$$

But this follows from Fubini-Tonelli.

> Note that if (3) holds for Borel sets, then (3) holds for Lebesgue null sets.

Suppose now that f is just Lebesgue measurable, and let G be open in $\mathbb{R}$.

Then $f^{-1}(G) = H \bigcup N$ where $H \in G_\delta$ and $m(N) = 0$ and

$$T^{-1}(f^{-1}(G)) = T^{-1}(H) \bigcup T^{-1}(N).$$

But by the first part, $T^{-1}(H)$ is still Borel, and $T^{-1}(N)$ is still null. So $f \circ T$ is measurable.

> Note that this kind of thing usually works – just establish something Borel sets, then use this characterization to extend it to Lebesgue.

22 Tuesday October 22

22.1 Convolution

Recall:

- **Continuous Compact Approximation:** $C_c \hookrightarrow L^1$ is dense.
- **Continuity in L^1:**

$$f \in L^1 \implies \|\tau f - f\| \to 0, \text{ i.e. } \lim_{y \to 0} \int |f(x+y) - f(x)| \, dx = 0.$$

- If $f \in L^1$, then for any $\varepsilon > 0$,
  - **Small tails:** There exists a δ such that for all E such that

$$m(E) \leq \delta \implies \int_E |f| < \varepsilon.$$

  - There exists an N such that

$$\int_{\{\|x\| \geq N\}} |f| < \varepsilon.$$

* Note that $|f(x)| < \varepsilon \ \forall x$ such that $\|x\| \geq N$ exactly when f is uniformly continuous

Definition: The *convolution* of f, g measurable functions on $\mathbb{R}^n$ is given by

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

for every x for which this integral makes sense.

Remarks:

- There are sufficient conditions on f, g which guarantee that $f \star g$ exists.
- If for some x, the function

  $$y \mapsto f(x-y)g(y)$$

  is measurable, then the function

  $$y \mapsto f(y)g(x-y)$$

  is also integrable.

> Note that this is just a translation followed by a reflection, which is still integrable since this operation is in $\mathrm{GL}(n, \mathbb{R})$) and $f \star g = g \star f$.

22.2 Properties of Convolutions

Theorem 1:

a.

$$f \in L^1 \text{ and } g \text{ bounded} \implies f \star g \quad \text{is bounded *and* uniformly continuous.}$$

b.

$$f, g \in L^1 \text{ and } f, g \text{ bounded} \implies \lim_{|x| \to \infty} (f \star g)(x) = 0.$$

> Note that (b) immediately follows if it were the case that $f \star g$ were uniformly continuous and *integrable*, but we don't necessarily need integrability for this result.

> Note: It is possible to pointwise multiply 2 integrable functions and get something non-integrable – consider f^2 where
>
> $$f(x) = \frac{1}{\sqrt{x}} \chi_{[0,1]}.$$

Theorem 2:

$$f, g \in L^1 \implies \|f \star g\|_1 \leq \|f\|_1 \|g\|_1,$$

and equality is attained if $f, g \geq 0$.

That is,

$$\int |f \star g| \leq \int |f| \int |g|.$$

Corollary: If g is additionally *bounded*, then

$$\lim_{|x| \to \infty} f \star g(x) = 0.$$

Theorem 3:

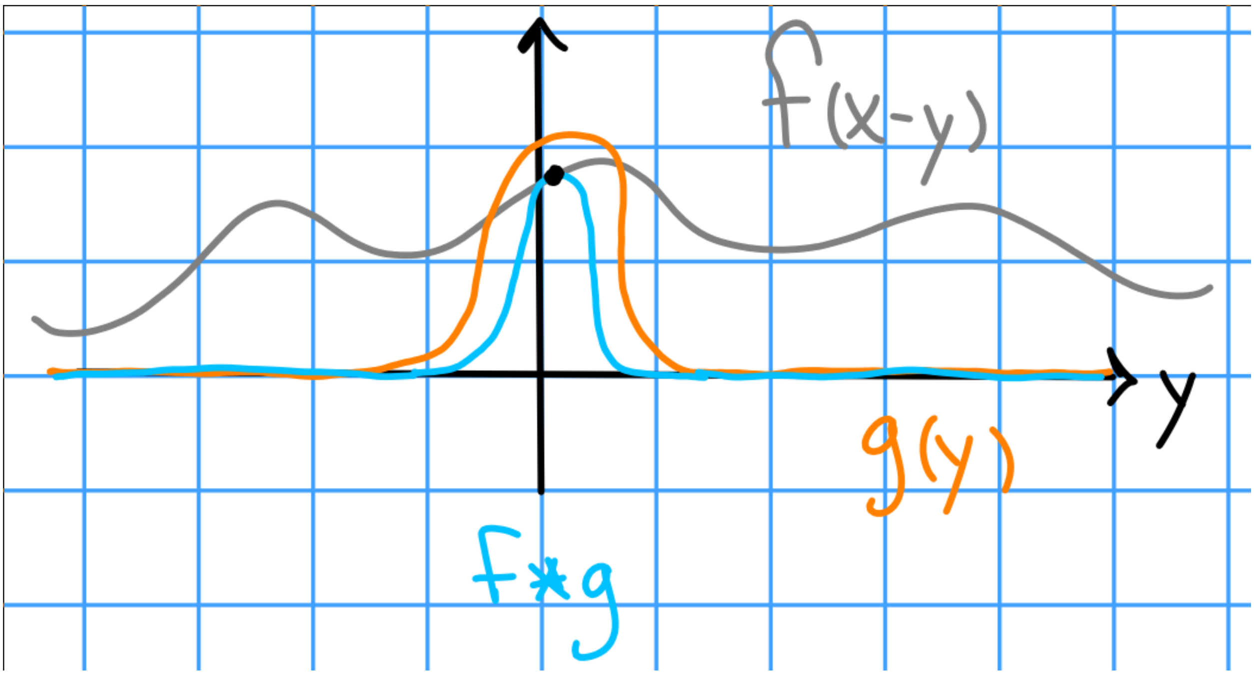
$$f \in L^1, \quad g \text{ differentiable}, \quad \text{and } g, \frac{\partial g}{\partial x_1}, \cdots, \frac{\partial g}{\partial x_n} \text{ all bounded} \implies$$
$$f \star g \in C^1 \text{ and } \frac{\partial}{\partial x_j}(f \star g) = f \star (\frac{\partial}{\partial x_j} g).$$

Corollary:

$$f \in L^1 \text{ and } g \in C_c^\infty \implies f \star g \in C^\infty \text{ and } \lim_{|x| \to \infty} f \star g(x) = 0.$$

In other words, defining C_0 as the functions that vanish at infinity, we have $f \star g \in C_0^\infty$.

Note that we don't necessarily preserve *compact support* after this convolution. See the following picture, which looks similar for any fixed x – particularly any large x.



Proof of Theorem 1, part (a):

$$
\begin{aligned}
\left|\int f(x-y)g(y)\ dy\right| &\leq \int |f(x-y)||g(y)|\ dy \\
&\leq M \int |f(x-y)|\ dy \\
&\leq M\|f\|_1.
\end{aligned}
$$

and

$$
\begin{aligned}
|f \star g(x+h) - f \star g| &= \left|\int f(x+h-y)g(y)\ dy - \int f(x-y)g(y)\ dy\right| \\
&\leq \int |f(x+h-y) - f(x-y)||g(y)|\ dy \\
&\leq M \int |f(z+h) - f(z)|\ dz \to 0.
\end{aligned}
$$

Proof of Theorem 1, part (b):

Let $\varepsilon > 0$, then choose N such that

$$
\int_{\{\|y\| \geq N\}} |f(y)|\ dy < \varepsilon \quad \text{and} \quad \int_{\{\|y\| \geq N\}} |g(y)|\ dy.
$$

Since $|x| \leq |x-y| + |y|$ by the triangle inequality, if we take $|x| \geq 2N$, then *either*

- $|x-y| \geq N$, or
- $|y| \geq N$.

In the first case, let $A_x = \{|x| \geq N\}$

$$
\begin{aligned}
|f \star g| &\leq \int |f(x-y)||g(y)|\ dy \\
&\leq M \int_{A_{x-y}} |f(x-y)| < M\varepsilon.
\end{aligned}
$$

and in the second case, take

$$
\begin{aligned}
|f \star g| &\leq \int |f(x-y)||g(y)|\ dy \\
&\leq M \int_{A_y} |g(y)| < M\varepsilon.
\end{aligned}
$$

Proof of Theorem 2:

Since $f, g \in L^1$, the function $h(x,y) \coloneqq f(x-y)g(y)$ will be measurable on $\mathbb{R}^n \times \mathbb{R}^n$ as a product of measurable functions if we can show that the function $f_{x,y} \coloneqq (x,y) \mapsto f(x-y)$ is measurable.

To see that this is the case, define $F(x-y,y) = f(x-y)$ by taking the cylinder, then let

$$T = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \implies T(x,y) = (x-y, y),$$

Thus $f_{x,y}(x,y) = (F \circ T)(x,y)$.

We can now note that

$$\begin{aligned}
\int\int |f(x-y)||g(y)| \; dy \; dx &=_{FT} \int\int |f(x-y)||g(y)| \; dx \; dy \\
&= \int |g(y)| \left(\int |f(x-y)| \; dx \right) \; dy \\
&= \|f\|_1 \|g\|_1.
\end{aligned}$$

This proves that the integrand is in $L^1(\mathbb{R}^{2n})$, so Fubini implies that $f \star g(x)$ is in L^1 for almost every x.

But then

$$\begin{aligned}
\int |f \star g(x)| \; dx &\leq \int\int |f(x-y)g(y)| \; dy \; dx \\
&= \|f\|_1 \|g\|_1.
\end{aligned}$$

■

Note that equality is attained here if $f, g \geq 0$.

23 Thursday October 24?

Todo.

24 Tuesday October 29

24.1 Approximations of the Identity

Theorem: Let $\phi \in L^1$ and $\displaystyle\int \phi = 1$.

Then

- If f is bounded and uniformly continuous, then $f * \phi_t \xrightarrow{u} f$ uniformly where
$$\phi_t(x) := \frac{1}{t^n}\phi(\frac{x}{t}).$$
- If $f \in L^1$, then $f * \phi_t \xrightarrow{L^1} f$ in L_1.

Applications:

24.2 Theorem 1: Smooth Compactly Supported Functions are Dense in L^1

Theorem: $C_c^\infty \hookrightarrow L^1$ is dense,

That is, $\forall \varepsilon > 0$ and for all $f \in L^1$, there exists a $g \in C_c^\infty$ such that $\|f - g\|_1 < \varepsilon$.

Proof: Since C_c^0 is dense in L^1, it suffices to show the following:

$$\forall \varepsilon > 0 \ \&\ h \in C_c^1, \quad \exists g \in C_c^\infty \ \Big|\ \|h - g\|_1 < \varepsilon.$$

Let $\phi \in C_c^\infty$ be arbitrary where $\displaystyle\int \phi = 1$ *(which exist!)*.

Then $\|h * \phi_t - h\|_1 < \varepsilon$ for t small enough. It remains to show that $f := h * \phi_t \in C_c^\infty$.

f is smooth because of theorem 3 regarding convolution, applied infinitely many times.

f is also compactly supported: since h, ϕ_t are compactly supported, so there is some large N such that $|x| > N \implies h(x) = \phi_t(x) = 0$.

Then if $|x| > 2N$, we can note that

$$|x| \leq |x + y| + |y|,$$

so either $|x - y| \geq 2N$ or $|y| \geq N$.

But then

$$f(x) := h * \phi_t(x) = \int h(x - y)\phi_t(y)\ dy = 0,$$

where by the previous statement, at least one term in the integrand is zero and thus the integral is zero and $f := h * \phi_t$ compactly supported.

∎

24.3 Theorem 2: Weierstrass Approximation:

Theorem: A function can be *uniformly* approximated by a polynomial on any closed interval, i.e.

$$\forall \varepsilon > 0,\ f \in C([a, b]), \quad \exists \text{ a polynomial } P \ \Big|\ |f(x) - P(x)| < \varepsilon \quad \forall x \in [a, b].$$

Proof: Let g be a continuous function on $[-M, M] \supseteq [a, b]$ such that $g\big|_{[a,b]} = f$.

Let $\phi(x) = e^{-\pi x^2}$ be the standard Gaussian, then $g * \phi_t \rightrightarrows g$ on $[-M, M]$, and thus $g * \phi_t \rightrightarrows f$ on $[a, b]$.

> The problem is that this is not a polynomial.

We can let $\varepsilon > 0$, then there is a t such that

$$|g * \phi_t(x) - g(x)| < \varepsilon \quad \forall x \in [-M, M].$$

Note that $\phi_t(x) = \dfrac{1}{t}e^{-\pi x^2/t^2}$, and Maclaurin expand to obtain

$$P(t) := \frac{1}{t}\sum_{n=0}^{\infty} \frac{(-1)^n \pi^n x^{2n}}{t^{2n} n!}.$$

> Note that the Maclaurin series will converge uniformly on compact sets!

By uniform convergence of P, we can truncate it to bound the difference by say $\varepsilon / \|g\|_1$.

Let $Q(x)$ be the truncated series. Then

$$|g * \phi_t(x) - g * Q(x)| \leq |g * (\phi_t - Q)(x)| \leq \|g\| \|p_t(x) - Q(x)\|_\infty < \varepsilon \to 0,$$

where $\|f\|_\infty = \sup_{x \in [a,b]} |f(x)|$ and $(g * Q)(x)$ is a polynomial. ■

24.4 Fourier Transform on $\mathbb{R}^n$

Given $f \in L^1$, we defined the Fourier transform of f by

$$\hat{f}(\xi) = \int f(x) e^{-2\pi i x \cdot \xi} \, dx.$$

Some facts we know about the Fourier transform:

- $f \in L^1 \implies \hat{f}$ is bounded and uniformly continuous.

  (From an old homework!)
- The Riemann-Lebesgue lemma: $\lim_{|\xi| \to \infty} \hat{f}(\xi) = 0$, i.e. $\hat{f}$ vanishes at infinity.

> Warning: it is **not** true that $f \in L^1 \implies \hat{f} \in L^1$!

24.4.1 Fourier Inversion Formula

Theorem (Inversion Formula): If $f, \hat{f} \in L^1$ then

$$f(x) = \int \hat{f}(x) e^{2\pi i x \cdot \xi} \, d\xi \quad \text{for a.e. } x,$$

i.e. $\hat{\hat{f}} = f(-x)$, and the Fourier transform is 4-periodic.

> Note that there is an interpretation here as writing an arbitrary function as a (continuous) sum of *characters*, where we're considering $\mathbb{R}^n$ with the action of translation. In this setting, exponentials are certain eigenfunctions.

Corollaries:

1. $f, \hat{f} \in L^1$ implies that f itself is bounded, continuous, and vanishes at infinity. (Note that this is not true for arbitrary L^1 functions!)

   > We will in fact show that $\left\{ f \;\middle|\; f, \hat{f} \in L^1 \right\} \hookrightarrow L^1$ is dense.

2. $f \in L^1$ and $\hat{f} = 0$ a.e. $\implies$ $f = 0$ almost everywhere

   (Proof uses the Inversion formula)

Proof of Inversion Formula:

Note: Fubini-Tonelli won't work here *directly*.

We'll have

$$f(x) = \int \int f(y) e^{-2\pi i y \cdot \xi} e^{2\pi i x \cdot \xi} \, dy \, d\xi,$$

which is (obviously?) not in $L^1(\mathbb{R}^{2n})$.

So we'll introduce a "convergence factor" $e^{-\pi t^2 |\xi|^2}$, which will make the integral swap result in something integrable, then take limits.

Important example (HW):

$$g(x) = e^{-\pi |x|^2} \implies \hat{g}(\xi) = e^{-\pi |\xi|^2}.$$

Note that

$$g_t(x) = \frac{1}{t^n} e^{-\pi |x|^2 / t^2}$$

is an approximation to the identity, and $\displaystyle\int g_t = 1$.

By a HW exercise, have have

$$\hat{g_t}(\xi) = \hat{g}(t\xi) = e^{-\pi t^2 |\xi|^2},$$

which is exactly the convergence factor we're looking for. Moreover, $f * g_t \xrightarrow{L^1} f$ in L^1.

This says that the Fourier transform "commutes with dilation" in a certain way.

Lemma (Multiplication Formula): If $f, g \in L^1$, then an easy application of Fubini-Tonelli yields

$$\int f \hat{g} = \int \hat{f} g.$$

We have

$$\begin{aligned}
\int \hat{f}(\xi) e^{-\pi t^2 |\xi|^2} e^{2\pi i x \cdot \xi} \, d\xi &:= \int \hat{f}(\xi) \phi(\xi) \qquad (= f * g_t(x) \xrightarrow{L_1} f) \\
&= \int f(y) \hat{\phi}(y) \, dy \\
&=_{DCT} \int \hat{f}(\xi) e^{2\pi i x \cdot \xi} \, d\xi \quad \text{as } t \to 0.
\end{aligned}$$

where $\phi(\xi) = e^{2\pi i x \cdot \xi} \hat{g_t}(\xi)$.

By a HW problem, we know

$$\hat{\phi}(y) = \hat{g}(y - x) = g_t(x - y).$$

But now one term is converging to $\displaystyle\int \hat{f}(\xi) e2\pi i x \cdot \xi \, d\xi$ as $t \to 0$ pointwise, and $f * g_t(x) \to f$ as $t \to 0$ in L_1.

So there is a subsequence of the latter term converging to f almost everywhere, and thus the pointwise limit in the first is equal to the L^1 limit in the second.

We thus obtain

$$f(x) = \int \hat{f}(\xi) e^{2\pi i x \cdot \xi} \, d\xi$$

almost everywhere.

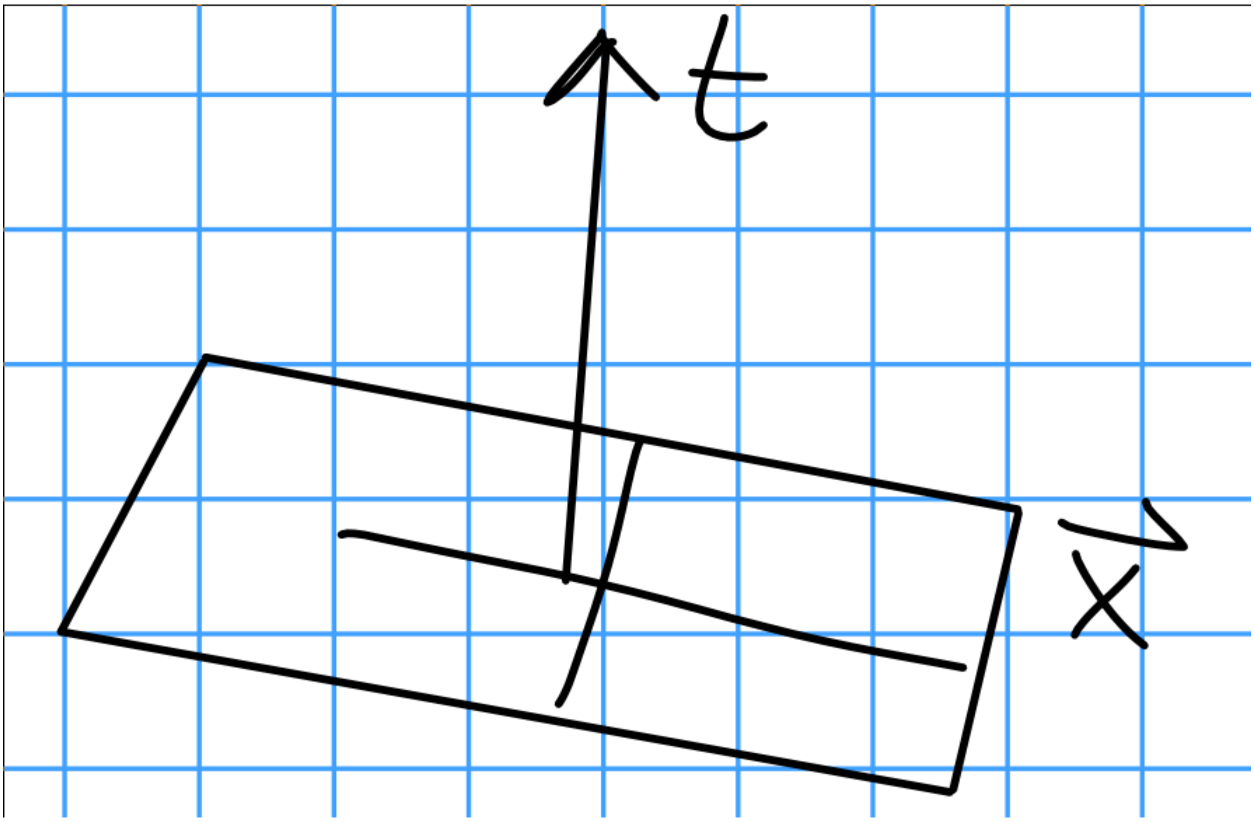
$\blacksquare$

25 Thursday October 31

Today: Some topics in PDEs.

25.1 The Heat / Diffusion Equation in the Plane

Setup: Let $\in \mathbb{R}^2$ be a plate, and consider it evolving over time t.



So we have pairs $(x, t) \in \mathbb{R}^2 \times \mathbb{R}_{\geq 0}$.

We have some initial distribution of heat on the plate, we want to know how it evolves over time.

This is modeled by the equation

$$\begin{aligned} \frac{\partial u}{\partial t} = \frac{1}{4\pi} \left(\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} \right) &\coloneqq \frac{1}{4\pi} \Delta u \\ u(x, 0) &= f(x). \end{aligned}$$

Consider a point and a small ball around that point. Then heat flow at any point x_0 is given by $\nabla_x u(x_0, t)$. Now think about the change in energy contained in this ball. We should have

$$\begin{aligned} \frac{\partial}{\partial t} \int_B u(x, t)\ dx &= \text{Flux across boundary} \\ &= \int_B \nabla \cdot \nabla_x u(x, t)\ dx \quad \text{by Green's/Divergence theorem} \\ &\coloneqq \int_B \Delta_x u(x, t)\ dx, . \end{aligned}$$

which is the heat equation.

25.1.1 Solution

We can use Fourier transforms to help solve these. Recall the identities:

- $\widehat{\frac{\partial}{\partial x_j} f}(\xi) = 2\pi i \xi_j \widehat{f}(\xi)$.
- $\widehat{\frac{\partial^2}{\partial x_j^2} f}(\xi) = (2\pi i \xi_j)^2 \widehat{f}(\xi) = -4\pi^2 \xi^2 \hat{f}(\xi)$.
- $\widehat{\Delta f}(\xi) = 4\pi^2 |\xi|^2 \hat{f}(\xi)$.

If we take the Fourier transform in the x variable, we get

$$\widehat{\frac{\partial u}{\partial t}} = \frac{\partial}{\partial t} \hat{u}(\xi, t) = -\pi |\xi|^2 \hat{u}(\xi, t).$$

Then the boundary conditions become $\hat{u}(\xi, 0) = \xi f(\xi)$. But note that this is now a first order ODE!

This is easy to solve, we get

$$\hat{u}(\xi, t) = c(\xi) e^{-\pi |\xi|^2 t} = \hat{f}(\xi) e^{-\pi |\xi|^2 t}.$$

But then

$$e^{-\pi |\xi|^2 t} = \hat{G}(\sqrt{t} \xi)$$

where $G(x) = e^{-\pi|x|^2}$.

We now have $\hat{u} = \hat{f}\hat{G} = \widehat{f * G}$, but **if the transforms are equal then the original functions are equal by the inversion formula**.

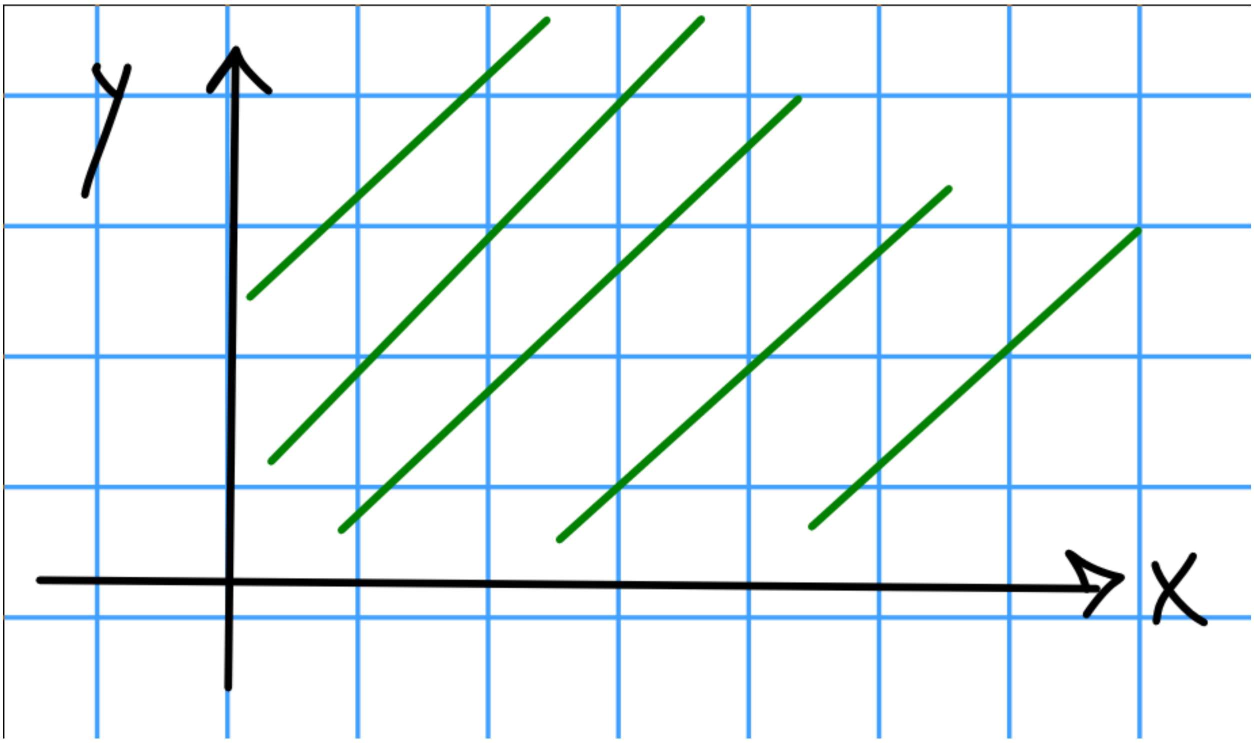
We thus obtain

$$u(x,t) = f * G_{\sqrt{t}}(x) \text{ where } G_{\sqrt{t}}(x) = \frac{1}{t^{n/2}} e^{-\pi|x|^2/t}.$$

Note that $f * g \to f$ as $t \to 0$, which matches with the original boundary conditions, and $f * g \to 0$ as $t \to \infty$, which corresponds with heat dissipating.

25.2 Dirichlet problem in the upper half-plane

Setup:



We want to solve

$$\Delta u = 0$$
$$u(x,0) = f(x).$$

25.2.1 Solution

We'll use the same technique as the heat equation, and obtain

$$\Delta u = 0 \implies -4\pi^2|\xi|^2\hat{u}(\xi, y) + \frac{\partial^2}{\partial y^2}\hat{u}(\xi, y) = 0$$

But this is a homogeneous second order ODE, so we can look at the auxiliary polynomial. If we have distinct roots, the general solution is $c_1 e^{r_1 x} + c_2 e^{r_2 x}$.

We thus obtain

$$\hat{u}(\xi, y) = A(\xi)e^{-2\pi|\xi|y} + B(\xi)e^{2\pi|\xi|y}$$

In particular, we can just take the first term, since the second term won't vanish at infinity. We again find that $A(\xi) = \hat{f}(\xi)$ by checking initial conditions, so

$$\hat{u}(\xi, y) = \hat{f}(\xi)\hat{P}(y\xi) = \widehat{f * P_y} \qquad \text{where } P(x) = \frac{1}{\pi}\frac{1}{1+x^2},$$

and $f * P_y \to f$ as $y \to 0$ as desired. . ## Wave Equation (Cauchy Problem in $\mathbb{R}^n$)

Same situation as the heat equation, but now in $\mathbb{R}^n \times \mathbb{R}_{\geq 0}$:

$$\begin{aligned}
\frac{\partial^2 u}{\partial t^2} &= \Delta_x u \\
u(x, 0) &= f(x) \\
\frac{\partial u}{\partial t}(x, 0) &= g(x).
\end{aligned}$$

This models something like plucking a string with initial shape f and initial velocity g.

> Note that this now involves a *second* derivative!

25.2.2 Solution

Using the same technique, we have

$$\begin{aligned}
\frac{\partial^2}{\partial t^2}\hat{u}(\xi, t) &= -4\pi^2|\xi|^2\hat{u}(\xi, t) \\
\hat{u}(\xi, 0) &= \hat{f}(\xi) \\
\frac{\partial}{\partial t}\hat{u}(\xi, 0) &= \hat{g}(\xi).
\end{aligned}$$

This is again 2nd order linear homogeneous, except there is now a complex conjugate pair of roots, so we get

$$\hat{u}(\xi, t) = \hat{f}(\xi)\cos(2\pi|\xi|t) + \frac{\hat{g}(\xi)\sin(2\pi|\xi|t)}{2\pi|\xi|}.$$

Note that the derivative of the first term is exactly the second term, so we have

$$u(x,t) = f * \frac{\partial}{\partial t} W_t(x) + g * W_t(x), \quad \hat{W}_t(\xi) = \frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}.$$

From the homework problems, we know:

- $n = 1$ implies $\chi_{[-1,1]}(x)$
- $n = 2$ implies $\frac{1}{\sqrt{1-|x|^2}}\chi_{-1,1}(x)$
- $n = 3$ implies we only get a measure, i.e. $w(x) = \sigma(x)$ where σ is a surface measure on S^2.
- For $n > 3$, W is a *distribution.*

Note that there is a solution given by D'Alembert,

$$u(x,t) = \frac{1}{2}(f(x+t) + f(x-t)) + \frac{1}{2}\int_{x-t}^{x+t} g(y)\ dy$$

Note the similarities – the first term is a rough average, the second term is a more continuous average.

Exercise: Verify that these two solutions are equivalent.

26 Tuesday: November 5

26.1 Hilbert Spaces

> See notes on the webpage.

Definition: An *inner product* on a vector space satisfies

- $\langle ax + by,\ z\rangle = a\langle x,\ z\rangle + b\langle y,\ z\rangle$ i.e., for all fixed $z \in V$, the map $x \mapsto \langle x,\ z\rangle$ is a *linear functional.*
- $\langle x,\ y\rangle = \overline{\langle y,\ x\rangle}$
- $\langle x,\ x\rangle \in (0, \infty)$

This induces a *norm*, $\|x\| = \langle x,\ x\rangle^{1/2}$.

Proposition 1: The map $x \mapsto \|x\|$ does in fact define a norm.

> The key to establishing this is the triangle inequality, since many of the other necessary properties fall out easily.

We'll need the **Cauchy-Schwarz inequality**, i.e.

$$|\langle x,\ y\rangle| \leq \|x\|\|y\|,$$

with equality iff $x = \lambda y$.

Note that this relates an inner product to a norm, as opposed to other inequalities which relates norms to other norms.

A useful computation:

$$\begin{aligned} \|x+y\|^2 &= \langle x+y,\ x+y\rangle \\ &= \|x\|^2 + 2\mathrm{Re}\langle x,\ y\rangle + \|y\|^2 \\ &\le \|x\|^2 + |\langle x,\ y\rangle| + \|y\|^2 \\ &\le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \qquad \text{by Schwarz} \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$

Definition: An inner product space that is *complete* with respect to $\|\cdot\|$ induced from its inner product is a *Hilbert space.*

Recall that a Banach space is a complete *normed* space.

Examples:

- $$\mathbb{C}^n \text{ with } \langle x,\ y\rangle = \sum x_j\overline{y_j}.$$

- $$\ell^2(\mathbb{N}) \text{ with } \langle x,\ y\rangle = \sum_{j=1}^{\infty} x_j y_j.$$

  Note that this is finite by **AMGM**, since by assumption

  $$\sum x_i y_i \le \frac{1}{2}(\sum x_i + \sum y_i) < \infty.$$

- $$L^2(\mathbb{R}^n) \text{ with } \langle f,\ g\rangle = \int f\overline{g}.$$

  This is also finite because

  $$\int |f\overline{g}| \le \frac{1}{2}(\int f + \int g).$$

Proof of Schwarz Inequality:

If $x = \lambda y$ for some $\lambda \in \mathbb{C}$, we have equality since

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

So we can assume $x - \lambda y \neq 0$ for *any* $\lambda \in \mathbb{C}$, so

$$\langle x - \lambda y,\ x - \lambda y\rangle > 0.$$

We thus have

$$0 < \langle x - \lambda y,\ x - \lambda y\rangle = \|x\|^2 - 2\overline{\lambda}\mathrm{Re}\langle x,\ y\rangle + |\lambda|^2\|y\|.$$

Now let $\lambda = tu$ where $t \in \mathbb{R}$ and $u = \langle x,\ y\rangle/|\langle x,\ y\rangle|$.

Then we get $0 < \|x\|^2 - 2t|\langle x,\ y\rangle| + t^2\|y\|^2$

But this is quadratic in t and doesn't have a real root, so its discriminant must be negative.

Thus

$$4|\langle x,\ y\rangle|^2 - 4\|y\|^2\|x\|^2 < 0,$$

which yields Cauchy-Schwarz.

■

26.2 Continuity of Norm and Inner Product

Application of the Schwarz Inequality:

If $x_n \to x$ in V, i.e. $\|x_n - x\| \to 0$, and similarly $y_n \to y$, we have $\langle x_n,\ y_n\rangle \to \langle x,\ y\rangle$ in $\mathbb{C}$.

Proof:

We have

$$\begin{aligned}
|\langle x_n,\ y_n\rangle - \langle x,\ y\rangle| &= |\langle x_n - x,\ y\rangle + \langle x,\ y_n - y\rangle| \\
&\leq |\langle x_n - x,\ y\rangle| + |\langle x,\ y_n - y\rangle| \\
&\leq \|x_n - x\|\|y\| + \|x\|\|y_n - y\| \qquad \text{by Schwarz} \\
&\to 0.
\end{aligned}$$

■

Exercise: Show $\|y_n - y\| \to 0$ iff $\|y_n\| \to \|y\|$.

Proposition (Parallelogram Law):

Let H be an inner product space, then

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

Exercise: Prove using parallelogram diagram.

Proof: Use the fact that

$$\|x \pm y\|^2 + \|x\|^2 \pm 2\mathrm{Re}\langle x,\ y\rangle + \|y\|^2,$$

so just add and the cross-terms will cancel.

Proposition (Pythagorean Theorem):

$$\langle x,\ y\rangle = 0 \implies \|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

In this situation, we say x, y are *orthogonal*.

Corollary: If $\{x_i\}$ are all pairwise orthogonal, then

$$\left\|\sum x_i\right\|^2 = \sum \|x_i\|^2.$$

26.3 Orthonormal Sets:

Definition: A countable collection $\{u_n\}$ is *orthonormal* iff

1. $\langle u_j,\ u_k\rangle = 0$ for $j \neq k$, and
2. $\langle u_j,\ u_j\rangle = \|u_j\|^2 = 1$ for all j.

Note: we only consider countable collections; a *separable* Hilbert space always has such a basis.

Definition: We say $\{u_n\}$ is an orthonormal *basis* for H if span $\{u_n\}$ (i.e. *finite* linear combinations of u_n) is dense in H.

26.4 Best Approximation and Bessel's Inequality

Theorem: Let $\{u_n\}$ be a countable orthonormal basis of H. Then for any $x \in H$, the *best approximation* to x by a sum $\sum_{n=1}^{N} a_n u_n$ when $a_n = \langle x,\ u_n\rangle$.

Note: these a_n will be Fourier coefficients later!

Proof:

$$\begin{aligned}
\left\|x - \sum a_n u_n\right\|^2 &= \|x\|^2 - 2\mathrm{Re}\sum \langle x,\ u_n\rangle a_n + \sum |a_n|^2 \\
&= \|x\|^2 - \sum |\langle x,\ u_n\rangle|^2 + \sum \left(|\langle x,\ u_n\rangle|^2 - 2\mathrm{Re}\langle x,\ u_n\rangle a_n + |a_n|^2\right) \\
&\leq \|x\|^2 - \sum |\langle x,\ u_n\rangle|^2 + |\langle x,\ u_n\rangle - a_n|^2 \qquad \geq 0,
\end{aligned}$$

where equality is attained iff $a_n = \langle x,\ u_n\rangle = a_n$. So this is the best approximation.

Note: Equalities are somehow easier to show – they necessarily involve direct computations.

But then

$$0 \leq \left\|x - \sum \langle x,\ u_n\rangle u_n\right\|^2 = \|x\|^2 - \sum |\langle x,\ u_n\rangle|^2,$$

so $\sum |\langle x,\ u_n\rangle|^2 \leq \|x\|^2$ holds for every N, and thus for the infinite sum, which is **Bessel's inequality**.

If this is an equality, then this is exactly **Parseval's theorem**.

26.5 Riesz-Fischer

Theorem (Riesz-Fischer):

1. The map

$$x \xrightarrow{\hat{}} \langle x,\ u_n \rangle \coloneqq \hat{x}(n)$$

   maps H onto ℓ^2 surjectively.

2. If $\{u_n\}^\infty$ is orthonormal in H and $\{a_n\}^\infty \in \ell^2(\mathbb{N})$, then there exists an $x \in H$ such that $\langle x,\ u_n \rangle = a_n$ for all $n \in N$.

3. x can be chosen such that $\|x\| = \sqrt{\sum |a_n|^2}$.

Remarks:

This is not a bijection: there may not be a unique such x. The a_n are referred to as the *Fourier coefficients*.

If

$$a_n = 0 \text{ for all } n \implies x = 0,$$

then the set $\{u_n\}$ is said to be **complete**.

This turns out to be equivalent to $\{u_n\}$ being a *basis*, which is equivalent to the convergence of Fourier series.

27 Thursday November 7

27.1 Bessel

Let H be a Hilbert space, then we have

Theorem (Bessel's inequality):

If $\{u_n\}$ is orthonormal in H, then for any $x \in H$ we have equation 0

$$\sum_n |\langle x,\ u_n \rangle|^2 \leq \|x\|^2,$$

or equivalently $\{\langle x,\ u_n \rangle\} \in \ell^2\mathbb{N}$.

Proof:

We have

$$0 \leq \left\| x - \sum_{n=1}^{N} \langle x,\ u_n \rangle u_n \right\|^2 = \|x\|^2 - \sum_{n=1}^{N} |\langle x,\ u_n \rangle|^2 \forall N. \tag{1}$$

Remark (Characterization of Basis):

TFAE:

- $$\operatorname{span}\{u_n\} = H,$$

  i.e. u_n is a basis.

- $$\sqrt{\sum_n |\langle x,\ u_n\rangle|^2} = \|x\| \forall x \in H,$$

  i.e. **Parseval's identity**

- $$\lim_{N\to\infty} \left\| x - \sum_n^N \langle x,\ u_n\rangle \right\| = 0,$$

  i.e. the Fourier series converges in H.

Recall the Riesz-Fischer theorem:

If $\{u\}_n$ is orthonormal in H and $\{a\}_n \in \ell^2(\mathbb{N})$, then

$$\exists x \in H \text{ such that } a_n = \langle x,\ u_n\rangle \text{ and } \|x\|^2 = \sum_n |a_n|^2.$$

Moreover, the map $x \mapsto \hat{x}(u) := \langle x,\ u_n\rangle$ maps H onto $\ell^2(\mathbb{N})$ surjectively.

Remark: This x is only unique if $\{u\}_n$ is *complete*, i.e. $\langle y,\ u_n\rangle = 0 \quad \forall n \implies y = 0$.

Proof: Let $S_N := \sum_{n=1}^N a_n u_n$.

Then S_N is Cauchy, so

$$\begin{aligned} \|S_N - S_M\|^2 &= \left\| \sum_{n=M+1}^N a_n u_n \right\|^2 \\ &= \sum_{n=M+1}^N \|a_n u_n\|^2 \qquad \text{by Pythagoras since the } u_n \text{ are orthogonal} \\ &= \sum_{n=M+1}^n |a_n| \to 0, \end{aligned}$$

since $\sum |a_n| < \infty$ implies that the sum is Cauchy.

Since H is complete, $S_N \to x$ for some $x \in H$.

We now need to argue that $a_n = \langle x,\ u_n\rangle$.

If $N \geq n$, then we have the identity

$$|\langle x,\ u_n\rangle - a_n| = |\langle x,\ u_n\rangle - \langle S_N,\ u_n\rangle| = |\langle x - S_N,\ u_n\rangle| \leq \|x - S_N\| \to 0.$$

Note: should be able to translate this to statements about epsilons almost immediately!

But then equation 1 holds in the limit as $N \to \infty$, which establishes equation 0. ■

Proof of characterization of basis:

1 $\implies$ 2: Let $\varepsilon > 0, x \in H, \langle x,\ u_n\rangle = 0$ for all n. We will attempt to show that $\|x\| < \varepsilon$, so $x = 0$. By (1), there is a $y \in \operatorname{span}\{u_n\}$ such that $\|x - y\| < \varepsilon$. But then $\langle x,\ y\rangle = 0$, so

$$\|x\|^2 = \langle x,\ x\rangle = \langle x,\ x - y\rangle \leq \|x\|x - y \leq \varepsilon\|x\| \to 0.$$

■

Note: $\langle x,\ x\rangle = \langle x,\ x\rangle - \langle x,\ y\rangle = \langle x,\ x - y\rangle$ since $\langle x,\ y\rangle = 0$.

2 $\implies$ 3: By Bessel, we have $\{\langle x,\ u_n\rangle\} \in \ell^2\mathbb{N}$, and we know that its norm is bounded by $\|x\|$. By Riesz-Fischer, there exists a $y \in H$ such that $\langle y,\ u_n\rangle = \langle x,\ u_n\rangle$ and $\|y\| = \sqrt{\sum |\langle x,\ u_n\rangle|^2}$. By completeness, we get $x = y$. ■

27.2 Existence of Bases

- Every Hilbert space has an orthonormal basis (possibly uncountable)
- H *separable* Hilbert space $\iff$ H has a *countable* basis (separable = countable dense subset).

Some examples of orthonormal bases:

- $$\ell^2\mathbb{N}: \quad u_n(k) = \mathbf{e}_n := \begin{cases} 1 & n = k \\ 0 & \text{otherwise} \end{cases}$$

- $$L^2([0,1]): \quad e_n(x) := e^{2\pi i n x}.$$

  Normed: by Cauchy-Schwarz, but need to show it's complete. Can use the fact that L^1 is complete.

  Note that

  $$\langle f,\ e_n\rangle = \int_0^1 f(x) e^{-2\pi i n x}\, dx,$$

  which is exactly the Fourier coefficient.

27.3 L^2 is Complete (Sketch)

Sketch proof that $L^2([0,1])$ is complete:

Note that $L^2([0,1]) \subseteq L^1([0,1])$, since

$$f \in L^2 \implies \int_0^1 |f| \cdot 1 \, dx \leq \sqrt{\int_0^1 |f|^2}$$

by Cauchy-Schwarz. This also shows that $\|f\|_1 \leq \|f\|_2$.

Let f_n be Cauchy in L_2. Then f_n is Cauchy in L^1, and since L^1 is complete, there is a subsequence converging to f almost everywhere.

By Fatou,

$$\int \liminf_k \left| f_{n_j} - f_{n_k} \right|^2 \leq \liminf \int \left| f_{n_j} - f_{n_k} \right|.$$

But the LHS goes to $\int \left| f_{n_j} - f \right|$ and the RHS is $\left\| f_{n_j} - f_{n_k} \right\| \to 0$, so less than ε if j is big enough. So $f_{n_j} \xrightarrow{L^2} f$ in L^2 as $j \to \infty$, and thus $f_n \to f \in L^2$ as $n \to \infty$.

27.4 Unitary Maps

Definition: Let $U : H_1 \to H_2$ such that $\langle Ux, \ Uy \rangle = \langle x, \ y \rangle$, i.e. U preserves angles, and we say U is *unitary.*

Then $\|Ux\| = \|x\|$, i.e. U is an *isometry.*

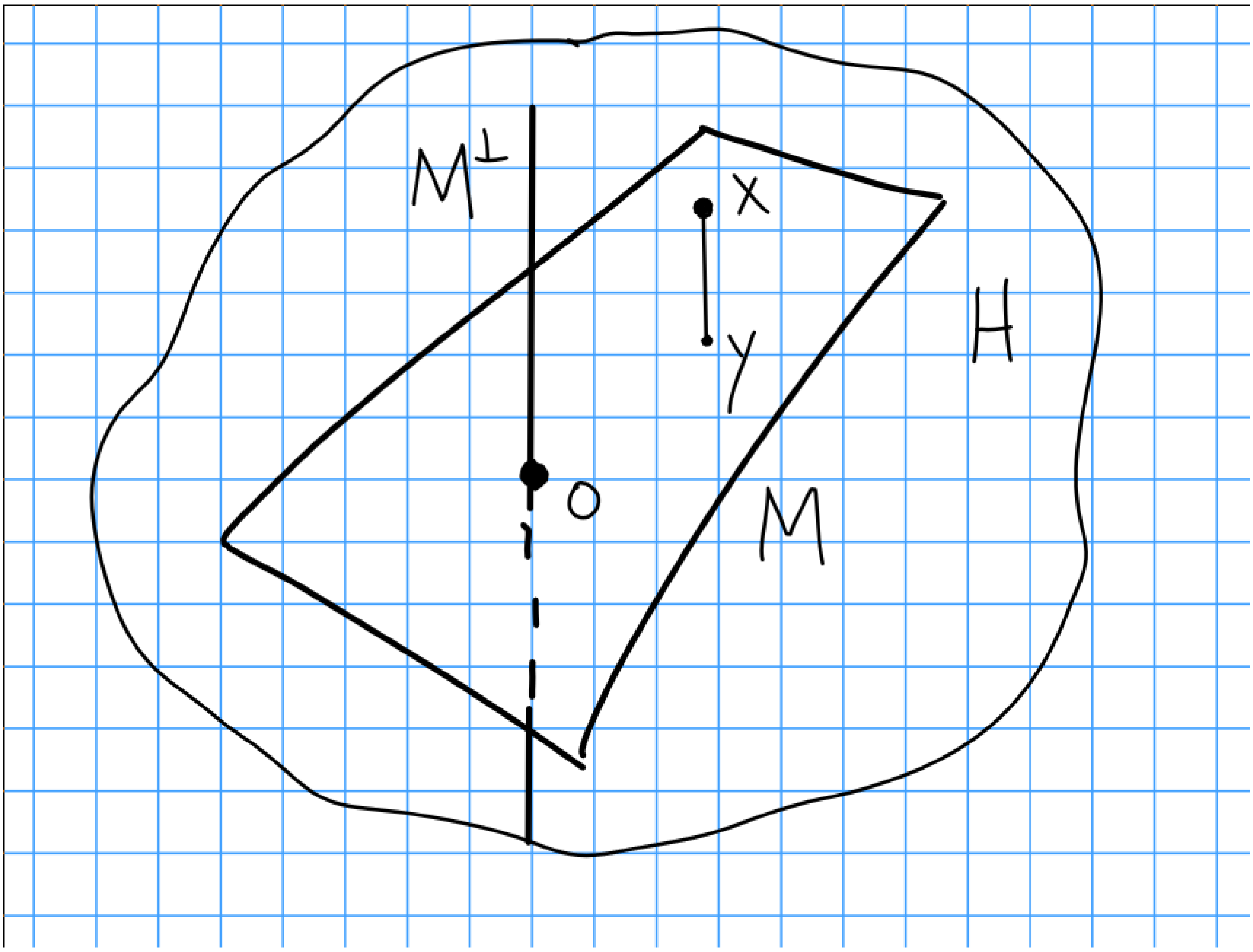
Every unitary map is an isometry. If U is surjective, this implication can be reversed.

For example, taking the Fourier transform yields

$$\sum \left| \hat{f}(u) \right|^2 = \|f\|_2^2 = \int |f|^2 \text{ and } \sum \hat{f}(u)\overline{\hat{g}(u)} = \int f\overline{g}.$$

A corollary of Riesz-Fischer: If $\{u\} - N$ is an orthonormal basis in H, then the map $x \mapsto \hat{x}(u) \coloneqq \langle x, \ u_n \rangle$ is a *unitary* map from H to ℓ^2.

So all Hilbert spaces are unitarily equivalent to $\ell^2\mathbb{N}$.



Subspaces in Hilbert spaces don't have to be closed, but orthogonal complements are always closed! See homework problem.

28 Tuesday November 12

28.1 Closed Subspaces and Orthogonal Projections

Definition: Let H be a Hilbert space, then a subspace $M \subseteq H$ is *closed* if $x_n \xrightarrow{H} x$ with $\{x_n\} \subset M$ implies that $x \in M$.

Note that finite-dimensional subspaces are *always* closed, so this is a purely infinite-dimensional phenomenon.

Proposition: Given any *set* M, then

$$M^{\perp} := \left\{ x \in H \,\middle|\, \langle x,\ y \rangle = 0 \ \forall y \in M \right\}$$

is always a closed subspace.

Proof: Homework problem.

Lemma: Let M be a closed subspace of H and $x \in H$. Then

1. There exists a unique $y \in M$ that is *closest* to y, i.e.

$$\exists y \in M \;\Big|\; \|x - y\| = \inf_{y' \in M} \|x - y'\|.$$

2. Defining $z \coloneqq x - y$, then $z \in M^{\perp}$.

Consequence 1: If $M \subseteq H$ is a closed subspace, then $(M^{\perp})^{\perp} = M$.

Note that $M \subseteq M^{\perp\perp}$ by definition. (Easy to check)

To show that $M^{\perp\perp} \subseteq M$, let $x \in M^{\perp\perp}$, then $x = y + z$ where $y \in M$ and $z \in M^{\perp}$.

Then

$$\langle x,\ z\rangle = \langle y,\ z\rangle + \langle z,\ z\rangle \implies \|z\|^2 = 0 \implies z = 0 \implies x = y.$$

Consequence 2:

Theorem: If $M \subseteq H$ is a closed subspace, then $H = M \oplus M^{\perp}$, i.e.

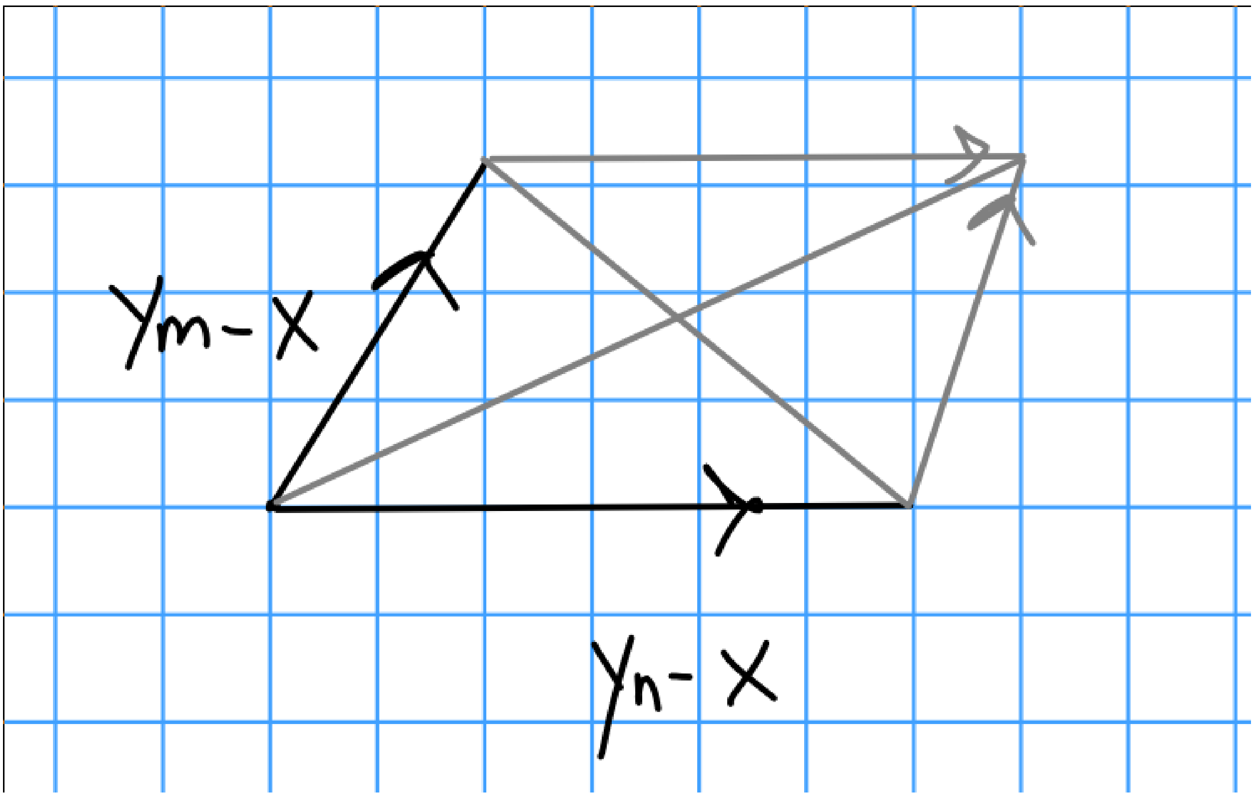
$$x \in H \implies x = y + z, \quad y \in M,\ z \in M^{\perp},$$

and y, z are the unique elements in $M, M^{\perp}$ that are closest to x.

Proof of Lemma (Part 1):

Let $\delta \coloneqq \inf_{y' \in M} \|x - y'\|$, which is a sequence of real numbers that is bounded below, and thus this infimum is attained. Then there is a sequence $\{y_n\} \subseteq M$ such that $\|x - y_n\| \to \delta$.

Consider the following parallelogram:



Then by the parallelogram theorem, we have

$$2(\|y_n - x\|^2 + \|y_m - x\|^2) = \|y_n - y_m\|^2 + \|y_n + y_m - 2x\|^2.$$

which yields

$$\begin{aligned}\|y_n - y_m\|^2 &= 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\left\|\frac{1}{2}(y_n + y_m) - x\right\|^2 \\ &\leq 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\delta^2 \to 0,\end{aligned}$$

since $\|y_n - x\|_H \to 0$ since $y_n \to_H x$.

It follows that $\{y_n\}$ is Cauchy in H, so $y_n \xrightarrow{H} y \in H$. But since the y_n were in M and M is closed, we in fact have $y \in M$. Since $\|x - y_n\| \to \|x - y\| = \delta$, we have the existence of x.

We'll establish uniqueness after part 2.

Proof of Lemma (Part 2):

Let $u \in M$, we want to show that

$$\langle z,\ u\rangle = \langle x - y,\ u\rangle = 0.$$

Without loss of generality, we can assume that $\langle z,\ u\rangle \in \mathbb{R}$, since u satisfies this property iff any complex scalar multiple does.

Let $f(t) = \|z + tu\|^2$ where $t \in \mathbb{R}$. Then $f(t) = \|z\|^2 + zt\langle z,\ y\rangle = t^2\|u\|^2$.

We know that t attains a minimum at $t = 0$, since $z + tu = x - (y + u)$, but y was the closest element to x and thus the norm is minimized exactly when $z + tu = x - y \implies t = 0$.

Because of this fact, we know that $f'(0) = 0$. But by using Calculus, we can compute that $f'(0) = 2\langle z,\ u\rangle$, so $\langle z,\ u\rangle$ must equal zero.

Now to show uniqueness, let $y' \in M$ and suppose $y' \neq u$ but $\|x - y'\| = \delta$. Then $x - y' = (x - y) + (y - y')$.

But these are two orthogonal terms, so we can apply Pythagoras to obtain

$$\begin{aligned}\|x - y'\|^2 &= \|x - y\|^2 + \|y - y'\|^2 \\ &\implies \delta = \delta + c \\ &\implies c = 0 \\ \implies &\|y - y'\| = 0 \\ &\implies y = y'.\end{aligned}$$

Note: the statement is the important thing here, less so this particular proof.

28.2 Trigonometric Series

Theorem: Let $e_n(x) := e^{2\pi i n x}$ for all $x \in [0,1]$ and $n \in \mathbb{Z}$. Then $\{e_n\}_{n\in\mathbb{Z}}$ is an *orthonormal basis* for $L^2([0,1])$.

Note: Orthonormality is easily check, so the crux of the proof is showing it's a basis.

Note: Elements in $\operatorname{span}\{e_n\}$ are referred to as *trigonometric polynomials.*

Goal: We'll show that the span of the trigonometric polynomials are dense in $L^2([0,1])$.

This will be a consequence of the following theorem:

28.2.1 Trigonometric Polynomials are Dense in $C^0([0,1])$

Theorem (Periodic Analogue of the Weierstrass Approximation Theorem): If $f \in C(\Pi)$ (where Π is a torus) and $\varepsilon > 0$, then there exists a trigonometric polynomial P such that $|f(x) - P(x)| < \varepsilon$ uniformly for all $x \in \Pi$.

Note that this measures closes in the *uniform* norm. We can relate these by

$$\|f(x) - P(x)\|_{L^2} \leq \|f(x) - P(x)\|_\infty, \text{ i.e. } \int_0^1 |f(x) - P(x)|^2 \leq \sup_x |f(x) - P(x)|^2.$$

Proof: Identify $\Pi = [-\frac{1}{2}, \frac{1}{2})$. Suppose there exists a sequence $\{Q_k\}$ of trigonometric polynomials such that

$$\begin{aligned} Q_k(x) &\geq 0 \quad \text{for all } x, k \\ \int_{-1/2}^{1/2} Q_k(x)\, dx &= 1 \quad \forall k \\ \forall \delta > 0, \quad Q_k(x) &\xrightarrow{u} 0 \text{ uniformly on } \Pi \setminus [-\delta, \delta]. \end{aligned}$$

Note that these properties are similar to what we wanted from approximations to the identity.

Define

$$P_k(x) = \int_{-1/2}^{1/2} f(y) Q_k(x-y)\, dy$$

by convolving over the circle, then P_k is also a trigonometric polynomial.

We then have

$$I = |P_k(x) - f(x)| \leq \int_{-1/2}^{1/2} |f(x-y) - f(x)| Q_k(y)\, dy \quad \text{by Property 2.}$$

We can now note that f is continuous on a compact set, so it is uniformly continuous, and thus for y small enough, we can find a δ such that

$$|f(x-y) - f(x)| < \varepsilon/2 \text{ for all } x \in B(\delta, x).$$

But this lets us break the integral into two pieces,

$$
\begin{aligned}
I &\leq \int_{y\in B_\delta} |f(x-y)-f(x)|Q_k(y)\ dy + \int_{y\in B_\delta^c} |f(x-y)-f(x)|Q_k(y)\ dy \\
&< \int_{y\in B_\delta} \frac{\varepsilon}{2}\ Q_k(y)\ dy + \int_{y\in B_\delta^c} |f(x-y)-f(x)|Q_k(y)\ dy \\
&\leq \int_{y\in B_\delta} \frac{\varepsilon}{2}\ Q_k(y)\ dy + \frac{\varepsilon}{2} \qquad \text{(for } k \text{ large enough)} \\
&\to 0 \qquad \text{since } Q_k \xrightarrow{u} 0.
\end{aligned}
$$

■

Constructing Q_k:

Define

$$
Q_k(x) = c_k \left(\frac{1+\cos(2\pi x)}{2}\right)^k,
$$

where c_k is chosen to normalize the integral to 1 to satisfy property 2. Property 1 is clear, so we just need to show 3,

Since $\cos(x)$ is decreasing on $[\delta, \frac{1}{2}]$,

$$
Q_k(x) \leq Q_k(\delta) = c_k \left(\frac{1+\cos(2\pi\delta)}{2}\right)^k.
$$

Note that the numerator is less than 2, so the entire term is a constant that is less than 1 being raised to the k power.

So this goes to zero exponentially, the question now depends on the growth of c_k. It turns out that $c_k \leq (k+1)\pi$, so it only grows linearly. So the whole quantity indeed goes to zero.

We can now write

$$
\begin{aligned}
1 &= 2c_k \int_0^{1/2} \left(\frac{1+\cos(2\pi x)}{2}\right)^k dx \\
&= 2c_k \int_0^{1/2} \left(\frac{1+\cos(2\pi x)}{2}\right)^k \sin(2\pi x) dx \\
&= \frac{2c_k}{\pi} \int_0^1 u^k\ du = \frac{2c_k}{\pi(k+1)}.
\end{aligned}
$$

■

Note: this is a nice proof!

Question: when is a function equal to its Fourier series? We have L^2 convergence, but when do we get pointwise?

Theorem from the 1960s: any L^2 function (in particular continuous functions) converges to its Fourier series *almost everywhere.*

29 Thursday November 14

Let $e_n(x) := e^{2\pi inx}$ for all $n \in \mathbb{Z}$ and $x \in [0,1]$.

Theorem: $\{e_n\}_{n\in\mathbb{Z}}$ is an orthonormal basis for $L^2([0,1])$.

> Note that $L_2([0,1]) = L^2(\Pi) = \left\{ f \in L^2 \;\middle|\; f(0) = f(1) \right\}$, since this only modifies a function at one point and we are identifying functions that agree almost everywhere.

29.1 Fourier Series

Definition: For any $f \in L^1(\Pi)$, we define its *Fourier coefficients*

$$\hat{f}(n) := \int_0^1 f(x)e^{-2\pi inx}\, dx \quad \forall n \in \mathbb{Z}$$

> Note that this resembles $\langle f,\ e_n \rangle$, although this is not an inner product space.

Definition: The *Fourier series* of f is defined as

$$\hat{f}(x) := \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{-2\pi inx}$$

> Note that this isn't necessarily *equal* to f, this only makes sense if the partial sums are converging to f in some sense.

Define the Nth partial sum

$$S_N f(x) := \sum_{|n|\leq N} \hat{f}(n)e^{2\pi inx}.$$

> Remark: We have $L_2(\Pi) \subseteq L^1(\Pi)$, so the Fourier coefficients *do* make sense here as an inner product for all $f \in L_2(\Pi)$.

Some consequences:

- By Riesz-Fischer, given any $\{a_n\} \in \ell^2(\mathbb{Z})$, there is a function $f \in L^2(\Pi)$ such that $\hat{f}(n) = a_n$ for all n.
- By Parseval,

$$\sum_{n\in\mathbb{Z}} \left|\hat{f}(n)\right|^2 = \int_0^1 |f(x)|^2\, dx$$

- $\lim_{N\to\infty} \|S_N f - f\|_2 = 0$, i.e. the Fourier series of f converges to f.

An answer: the Fourier series equals f in an L^2 sense, but recall that pointwise equality was a hard theorem proved only 50 or so years ago!

> Remark: note that for the Fourier transform, when $f \in L^1$ and $\hat{f} \in L^1$, we have
>
> $$f(x) = \int \hat{f}e^{-2\pi ix}\, dx,$$
>
> and we can get analogous statements here.

29.2 Uniform Convergence of Fourier Series

Theorem:

If $f \in L^1(\Pi)$ and $\left\{\hat{f}(n)\right\} \in \ell^1$, then $S_N f \xrightarrow{u} f$ *uniformly* on Π.

Corollaries:

1. If $f \in C^1(\Pi)$, then $S_N f \to f$ uniformly on Π.
2. $$f \in C(\Pi) \text{ and } f' \in L^2(\Pi) \implies S_N f \xrightarrow{u} f \text{ on } \Pi.$$

> Note that the first condition alone is not sufficient: there exists a continuous function whose Fourier series diverges at a point.

Proof: Exercise

> So if $f \in C^1$, it is *equal* to its Fourier series. Everyone should know this fact!

The following is a beautiful and fundamentally amazing fact:

Fact (The Riemann-Localization Principle):

If $f \in L^1(\Pi)$ and f is constant on some neighborhood of $x \in \Pi$, then $S_N f(x) \to f(x)$ pointwise at this particular x.

> Note that computing the Fourier coefficients requires integrating over the entire circle, but somehow the behavior of the function elsewhere doesn't matter at x!

Proof of theorem:

Since $\left\{\hat{f}(n)\right\} \in \ell^1(\mathbb{Z})$, this gives us what we need to apply the M test. So $S_N f \xrightarrow{u} g$ uniformly for some continuous g.

How can we argue $g = f$? Consider

$$\begin{aligned}
\hat{g}(n) &= \int_0^1 \left(\sum_{m \in \mathbb{Z}} \hat{f}(m) e^{2\pi i m x} \right) e^{-2\pi i n x} \, dx \\
&= \sum_{m \in \mathbb{Z}} \hat{f}(m) \int_0^1 e^{-2\pi i x(m-n)} \\
&= \sum_{m \in \mathbb{Z}} \hat{f}(m) \mathbb{1}\left[m = n\right] \\
&= \hat{f}(m),
\end{aligned}$$

so the question is now whether

$$\hat{g}(n) = \hat{f}(n) \ \forall n \implies g = f \text{ almost everywhere }.$$

> Note: if we accept this fact at face value, this proof only requires undergraduate analysis and uses facts about uniform convergence allowing sums to commute with integrals.

This will be true if $f \in \ell^2$. Why is this the case?

It's not strict inclusion, since

$$L^2([0,1]) \subseteq L^1([0,1]) \text{ but } \ell^1(\mathbb{Z}) \not\subseteq \ell^2(\mathbb{Z}).$$

We can use the fact that if $\sum |f_n| < \infty$, then $f_n \to 0$, and in particular f_n is bounded. So we have

$$\left(\sum |f_n|^2\right)^{1/2} \leq \sum |f_n| < \infty.$$

But then since $\ell^1 \subseteq \ell^2 \implies \left\{\hat{f}(n)\right\} \in \ell^2 \implies f \in L^2$ by Riesz-Fischer and completeness.

$\blacksquare$

An alternative proof: Again suppose $\left\{\hat{f}(n)\right\} \in \ell^1(\mathbb{Z})$, $S_N f \to g$.

Since $f \in L^2(\Pi)$ (because of the second assumption), we have $S_N f \to f$ in L^2.

> Use the fact that $f_n \to f \in L^2$ (or L^1) implies that a **subsequence converges almost everywhere.**

So $f = g$ almost everywhere. $\blacksquare$

29.3 Dual of a Vector Space

See notes on web page.

- Definition: Linear functional.
- Examples.
- Definition: Dual space.
  - Don't need X to be complete, just need a normed vector space.
- Theorem: A linear functional is continuous iff it is *bounded*.
- Definition: Bounded functional.
  - Can't define as $|L(x)| \leq C$ (since linearity allows pulling out scalars), so we can just restrict attention to $\left\{x \;\middle|\; \|x\| \leq 1\right\}$.
- $X^\vee$ is always a vector space
- $X^\vee$ is always a Banach space with norm $L \mapsto \sup\limits_{x \in X} |L(x)|$.
- $y \mapsto L_y := \langle \cdot\,,\ y \rangle$ is a conjugate linear isometry $H \to H^\vee$ which is surjective.
- Riesz Representation Theorem (for Hilbert spaces)
  - Use the fact that $L \in H^\vee \implies \ker L$ is a closed subspace.
  - $z \in M^\perp$, look at $\langle (Lx)z - (Lz)x,\ z \rangle$.

Upcoming:

- A bit about L^p spaces
- Dual of L^1, dual of L^∞.
- Abstract measure theory
- Hahn-Banach, Radon-Nikodym, and Lebesgue Density from the perspective of differentiation theorems.

30 Tuesday November 19

30.1 Lp Spaces

Given $f : \mathbb{R}^n \to \mathbb{C}$ and $0 < p < \infty$, we define

$$\|f\|_p = \left(\int |f|^p \right)^{1/p}.$$

and $L^p(\mathbb{R}^n) = \left\{ f \;\middle|\; \|f\|_p \infty \right\}$.

We also define

$$\|f\|_\infty = \inf_{a \geq 0} \left\{ m(\left\{ x \;\middle|\; f(x) > a \right\}) = 0 \right\}$$

which is morally the "best upper bound almost everywhere".

Qual problem alert: If $X \subseteq \mathbb{R}^n$ with $\mu(X) < \infty$ then $\|f\|_p \to \|f\|_\infty$.

Note that $|f(x)| \leq \|f\|_\infty$ almost everywhere, and if $|f(x)| \leq M$ almost everywhere, then $\|f\|_\infty \leq M$.

For $1 \leq p \leq \infty$, $(L^p, \|\cdot\|_p)$ yields a complete normed vector space. Scaling and non-degeneracy are fairly clear, it just remains to show the triangle inequality (sometimes referred to as Minkowski's inequality), i.e.

$$f, g \in L^p \implies \|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

For $p = 2$, this boiled down to Cauchy-Schwarz, here we'll need a souped-up version.

Definition: If $1 \leq p \leq \infty$, we define the *conjugate exponent* of p as the q satisfying $\dfrac{1}{p} + \dfrac{1}{q} = 1$.

An immediate consequence is that

$$q = \frac{p}{p-1}.$$

Holder's Inequality: If f, g are measurable functions then

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

Proof of Minkowski:

$$\begin{aligned}
|f+g|^p &= |f+g||f+g|^{p-1} \\
&\leq (|f|+|g|)|f+g|^{p-1} \\
&\implies \int |f+g|^p \leq \int |f||f+g|^{p-1} + \int |g||f+g|^{p-1} \\
&\leq \|f\|_p (\int |f+g|^{(p-1)q})^{1/q} + \|g\|_p (\int |f+g|^{(p-1)q})^{1/q} \\
&= (\|f\|_p + \|g\|_p) + (\int |f+g|^p)^{1-1/p} \\
&= (\|f\|_p + \|g\|_p) + (\int |f+g|^p)^{1/q},
\end{aligned}$$

and taking pth roots yields the result. (?? Revisit) ■

Note: For $1 \leq p \leq \infty$, L^p is a Banach space.

AM-GM: $\sqrt{ab} \leq \frac{a+b}{2}$.

Proof of Holder:

We'll use the following key fact:

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

with equality iff $a = b$. This can be verified by the first derivative test.

Important simplfication: we can assume that $\|f\|_p = \|g\|_q = 1$, since

$$\|fg\|_1 \leq \|f\|_p \|f\|_q \iff \int \frac{|f|}{\|f\|_p} \frac{|g|}{\|g\|_q}.$$

Applying the key fact, we can choose $\lambda = \frac{1}{p}, a = |f|^p, b = |g|^q$.

We then obtain

$$\int |f||g| \leq \int \frac{|f|^p}{p} \frac{|g|^q}{q} = \frac{1}{p} + \frac{1}{q} = 1.$$

■

30.2 Dual of L^p

Given $g \in L^q$, define an operation

$$\begin{aligned}
\Lambda_g(f) : L^p &\to \mathbb{C} \\
f &\mapsto \int fg.
\end{aligned}$$

Note that this makes sense by Holder's inequality, i.e. fg is integrable. This defines a linear functional on L^p, which is continuous by Holder, since we have

$$|\Lambda_g(f)| \leq \|g\|_q \|f\|_p$$

where $\|g\|_q$ is a constant that works for all $f \in L^p$.

> Recall: linear functionals are continuous iff bounded.

We have

$$\|\Lambda_g\|_{(L^p)^\vee} := \sup_{\|f\|_p = 1} \left| \int fg \right| \leq \|g\|_q.$$

> In fact, we have equality here for every $g \in L^q$. This is sometimes referred to as the converse of Holder.

Thus the map $g \mapsto \Lambda_g$ is an *isometric* map $L^q \hookrightarrow (L^p)^\vee$ for $1 \leq p, q \leq \infty$.

By Riesz representation, it turns out that this is a surjection as well for $p \neq \infty$.

Big Fact: This breaks for $p = \infty$, but for $1 \leq p < \infty$, this mapping is surjective.

30.3 Riesz Representation

Theorem (Riesz Representation): Suppose $1 \leq p < \infty$ and let q be its conjugate exponent, and let $X \subseteq \mathbb{R}^n$ be measurable.

Given any $\Lambda \in (L^p(X))^\vee$, there exists a unique $g \in L^q(X)$ such that for all $f \in L^p(X)$, we have

$$\Lambda(f) = \int_X fg \quad \text{and} \quad \|\Lambda\|_{(L^p(X))^\vee} = \|g\|_{L^q(X)}.$$

Summary:

- If $1 \leq p < \infty$, we have $(L^p)^\vee = L^q$.
- $(L^\infty)^\vee \supset L^1$, since the isometric mapping is always injective, but *never* surjective, so this containment is always proper (requires Hahn-Banach Theorem).

> For qual, supposed to know this for $p = 1$. $p = 2$ case is easy by Riesz Representation for Hilbert spaces.

Proof (in the special case where $1 \leq p < 2$ and $m(X) < \infty$):

We'll use the fact that we know this for $p = 2$ already.

Let $\Lambda \in (L^p)^\vee$, then we know

$$|\Lambda(f)| \leq \|\Lambda\|_{(L^p)^\vee} \|f\|_p,$$

since $\|\Lambda\|$ is the *best* upper bound.

> Note: in general, there are no inclusions between L_p, L_q, but restricting to a compact set changes this fact. Example from homework:
>
> $$L^2(X) \subseteq L^1(X) \text{ for } m(X) < \infty.$$

This follows from

$$\|f\|_1 \leq \|f\|_2 \|1\|_2 = m(X)^{1/2} \|f\|.$$

But this works for $L^2(X) \subseteq L^p(X)$ by taking

$$\|f\|_p^p = \int |f|^p \leq (\int |f|^2)^{p/2} (\int |1|^?)^{1-\frac{2}{p}}$$

by Holder with $\dfrac{2}{p}$.

So we can write

$$|\Lambda(f)| = \|\Lambda\| m(X)^{\frac{1}{p} - \frac{1}{2}} \|f\|_2 \quad \forall f \in L^2,$$

which verifies that Λ is a continuous linear functional on L^2, so $\Lambda \in (L^2)^\vee$, and by Riesz Representation, $\exists g \in L^2$ such that $\Lambda(f) = \displaystyle\int fg$ for all $f \in L^2$.

This is almost what we want, but we need $g \in L^q$ and $f \in L^p$. We also want to show that $\|\Lambda\| = \|g\|_q$.

Claim: $g \in L^q$ and $\|g\|_q \leq \|\Lambda\|$.

Pause on the proof, we'll come back to it!

Note that since $L_2 \subseteq L^p$ and both have simple functions as a dense subset, L^2 is in fact dense in L^p.

So let $f \in L^p$ and pick a sequence $f_n \subset L^2$ converging to f in the L^p norm.

Then $\Lambda(f_n) \to \Lambda(f)$ by continuity, and since $g \in L^q$, integrating against g is a linear functional $\Lambda_q(f_n)$ on L^q converging to $\displaystyle\int fg$, so $\Lambda(f) = \displaystyle\int fg$. ■

Definitely need to know: $(L^1)^\vee = L^\infty$!

Proof of claim: Suppose it's not true, so $\|g\|_\infty > \|\Lambda\|_{(L^1)^\vee}$.

Using the fact that $\|g\|$ is the best lower bound, there must be a positive measure set such that $|g(x)| \geq \|\Lambda\|$. So there is some set $E = \left\{ x \;\middle|\; |g(x) > \|\Lambda\|| \right\}$ with $m(E) > 0$.

Let

$$h = \frac{\overline{g}}{|g|} \frac{\chi_E}{m(E)}.$$

Note: useful technique!

Then $h \in L^2$ and $\|h\| = 1$.

Then

$$\Lambda(h) = \frac{1}{m(E)} \int_E |g| \geq \|\Lambda\| O(1),$$

which is a contradiction.

31 Thursday November 21

31.1 Abstract Measure Theory

Definition: Let X be a set and $\mathcal{M}$ be a σ-algebra of subsets of X.

Then $(X, \mathcal{M})$ is referred to as a *measurable* space, noting that we have not yet equipped it with a measure μ.

Definition: A *measure* μ on $(X, \mathcal{M})$ is a function $\mu : \mathcal{M} \to [0, \infty]$ such that

- The silly condition, $\mu(\emptyset) = 0$
- The important condition, $\mu(\coprod_{i \in \mathbb{N}} E_i) = \sum_{i \in \mathbb{N}} \mu(E_i)$.

Then $(X, \mathcal{M}, \mu)$ is called a *measure space*. Things can be measured in this setting, but more importantly, an integral can be defined.

Definition: A measure is σ*-finite* iff $X = \bigcup E_j$ with $m(E_j) < \infty$ for each j.

> Note: most measures encountered in practice seem to be σ-finite, so we could just as well incorporate this into our definition.

Examples:

- The Lebesgue measure
- Let $X = \{x_n\}_{n=1}^{\infty}$ a countable collection of objects, $\{\mu_n \in [0, \infty]\}$, and define $\mu(x_n) := \mu_n$.

  Then we can take the σ-algebra $\mathcal{M} = \mathcal{P}(X)$, so

$$\begin{aligned} \mu : \mathcal{P}(X) &\to [0, \infty] \\ E &\mapsto \sum_{\left\{ n \middle| \, x_n \in E \right\}} \mu_n. \end{aligned}$$

  In the special case $\mu_n = 1$ for all n, we have $\mu(E) = \#E$, the number of elements in E, which is the counting measure.

- Let $X = \mathbb{R}^n$ and let $\mathcal{M}$ be the Lebesgue measurable subsets, and let $\mu(E) = \int_E f$ for some fixed $f \in L^+$.

  > Exercise: Show that this defines a measure.

  In the special case $f \equiv 1$, we get the usual Lebesgue measure $\mu = m$. We write $d\mu := f dx$. Note that

$$m(E) = 0 \iff \mu(E) = 0,$$

  which is referred to as *absolute continuity.*

  > Note that all absolutely continuous measures occur in this way! But there are more exotic measures. Thinking about representability theorems, this says that measures are like "generalized integrable functions", but the collection of measures is richer.

- The Dirac mass:

$$\delta_0(E) = \begin{cases} 1 & 0 \in E \\ 0 & \text{else} \end{cases}.$$

31.2 Basic Properties of Measures

Fix a measure space $(X, \mathcal{M}, \mu)$.

1. **Monotonicity**:

$$E_1 \subseteq E_2 \implies \mu(E_1) \leq \mu(E_2).$$

   This follows from writing

$$E_2 = E_1 \coprod (E_2 \setminus E_1)$$

   and taking measures, which are always ≥ 0.

2. **Subadditivity**:

$$\mu(\bigcup E_i) \leq \sum \mu(E_i)$$

3. **Continuity from above and below:**

$$\begin{aligned} E_j \nearrow E &\implies \mu(E_j) \to \mu(E) \text{ and} \\ E_j \searrow E,\ \mu(E_1) < \infty &\implies \mu(E_j) \to \mu(E). \end{aligned}$$

Definition: A measure space is *complete* iff when $F \in \mathcal{M}$ is measurable, $\mu(F) = 0$, and $E \subseteq F$, we have $E \in \mathcal{M}$.

> Recall that the Lebesgue measure is complete, and the Borel measure is *not*. Review why this is the case!

31.3 Construction of Measures

Given an $(X, \mathcal{M})$, we construct μ in the following way:

1. Define an outer measure (or premeasure) μ^* on $\mathcal{P}(X)$.
2. Caratheodory:

$$E \subseteq X \text{ is measurable} \iff \mu_*(A) = \mu_*(A \bigcap E) + \mu_*(A \bigcap E^c) \quad \forall A \subseteq X.$$

> Note: it is worth recalling why this is equivalent to the usual "open set" definition, i.e. $\exists G$ open such that $\mu_*(G \setminus E < \varepsilon$, where we really needed a topology to talk about open sets.

3. Defining

$$\mathcal{M} := \{\text{Caratheodory measurable sets}\}$$

   yields a σ-algebra and $\mu_*|_{\mathcal{M}}$ is a measure.

31.4 Measurable Functions

Next up: define integrability, by first defining what it means for a function to be measurable.

Definition: A function $f : X \to \overline{\mathbb{R}}$ is *measurable* $\iff$ $\mu(\left\{x \in X \;\middle|\; f(x) > a\right\}) \in \mathcal{M}$ for all $a \in \mathbb{R}$.

We say two functions are *equal almost everywhere* if they disagree on a measure zero set, and we can define simple functions in a similar way.

Definition: If ϕ is simple, i.e. $\phi = \sum_{j=1}^{N} a_j \chi_{E_j} \in L^+$ (is non-negative), then

$$\int \phi \; d\mu \coloneqq \sum_j a_j \mu(E_j).$$

Then if $f \in L^+$, we define

$$\int f d\mu = \sup \left\{ \int \phi \; d\mu \;\middle|\; 0 \leq \phi \leq f, \phi \text{ is simple.} \right\}.$$

For f arbitrary and measurable, write $f = f_+ - f_-$, and define

$$\int f d\mu \coloneqq \int f_+ d\mu - \int f_- d\mu$$

whenever it makes sense (i.e. both are not infinite)

Consider an earlier example: given $(X, \mathcal{M}, \mu)$ and $f \in L^+(X, \mu)$, we can define

$$\mu_f(E) \coloneqq \int_E f d\mu \in \overline{\mathbb{R}}.$$

This always yields a measure, and moreover has the property $\mu(E) = 0 \implies \mu_f(E) = 0$.

> Note that we can actually generalize and let $f \in L^+$. Then the measure defined here can take on negative or even complex numbers, which turns out to be a useful (see "*signed measures*").

This is closely related to the usual notion of signed area between a curve and the x-axis we deal with in Calculus.

31.5 Absolute Continuity and Radon-Nikodym

Definition Let μ, ν be two measures on $(X, \mathcal{M})$. Then we say $\nu \ll \mu \iff \nu(E) = 0$ whenever $E \in \mathcal{M}$ and $\mu(E) = 0$, and that ν *is* absolutely continuous with respect to μ*.

Exercise: If ν is finite, i.e. $\nu(X) < \infty$, then

$$\nu \ll \mu \iff \forall \varepsilon > 0 \; \exists \delta > 0 \;\middle|\; \mu(E) < \delta \implies \nu(E) < \varepsilon,$$

which explains the terminology.

> Worth looking at more in-depth. Should be in textbook.

Theorem (Partial Radon-Nikodym): If μ, ν are two σ-finite measures on $(X, \mathcal{M})$ such that $\nu \ll \mu$, then there exists a unique non-negative function $f \in L^1(X, \mu)$ such that

$$d\nu = f d\mu \text{ and } \nu(E) = \int_E f d\mu \text{ for all } E \in \mathcal{M}.$$

Note: this is a representation theorem. This somehow all traces back to the Riesz Representation theorem for Hilbert spaces, which was a trivial proof! Worth recalling.

Proof (Sketch): We can assume μ, ν are σ-finite (there are standard techniques to do this).

Now define the measure $\rho : \nu + \mu$ and $L(\psi) = \int_X \psi d\nu$ for all $\psi \in L^2(X, \rho)$. Then L turns out to be a *continuous* linear functional on $L^2(X, \rho)$, which isn't completely obvious. This follows because it is bounded, since for all $\psi \in L^2(\rho)$ we have

$$\begin{aligned}\int |\psi| d\nu &\leq \int |\psi| d\rho \\ &\leq \|\psi\|_{L^2(\rho)} \rho(X)^{1/2} \\ &\leq C \|\psi\|_{L^2(\rho)},\end{aligned}$$

which follows from an application of Cauchy-Schwarz.

Then there exists a $g \in L^1(\rho)$ such that

$$\int \psi d\nu = \int \psi g d\rho = \int \psi g d\nu + \int \psi g d\mu.$$

By collecting terms, we obtain

$$\int_X \psi(1-g) d\nu = \int_X \psi g d\mu \quad \forall \psi \in L^2(\rho).$$

Now consider letting $\psi = \chi_E$ for some set. Then $\nu(E) = \int_E g d\rho$, from which it can be deduced that $0 \leq g \leq 1$ almost everywhere.

Since $\nu \ll \mu$, we actually have $0 \leq g < 1$ ρ-a.e. instead. This follows from taking $B = \{g(x) = 1\}$ and $\psi = \chi_B$ and using the above identity we found to deduce that $\mu(B) = 0$ and thus $\nu(B) = 0$ and $\rho(B) = 0$.

Now let

$$\psi = \chi_E(1) g + g^2 + \cdots g^n),$$

yielding

$$\begin{aligned}\int_E (1 - g^{n+1}) d\nu &= \int_E (1 + g + \cdots + g^n) d\mu \\ \to_{DCT} \nu(E) &= \int_E \frac{g}{1-g} d\mu,\end{aligned}$$

so we can take the integrand on the RHS to be our f.

■

Beautiful proof! Due to Von Neumann.

To show: the fundamental theorem of Calculus for measures, i.e. the Lebesgue differentiation theorem, which looks like

$$\lim_{r \to 0} \frac{1}{m(B_r(x))} \int_{B_r(x)} f(y)\ dy = f(x) \text{ a.e.}$$

and specializes when $f = \chi_E$.

32 Tuesday November 26th

32.1 Differentiation

Question: Let $f \in L^1([a,b])$ and $F(x) = \displaystyle\int_a^x f(y)\ dy$. Is F differentiable a.e. and $F' = f$?

If f is continuous, then absolutely yes.

Otherwise, we are considering

$$\frac{f(x+h) - F(x)}{h} = \frac{1}{h} \int_x^{x+h} f(y)\ dy \to_? f(x)$$

so the more general question is

$$\lim_{\substack{m(I) \to 0 \\ x \in I}} \frac{1}{m(I)} \int_I f(y)\ dy =_? f(x) \text{ a.e.}$$

Note that if f is continuous, since $[a,b]$ is compact, we have uniform continuity and

$$\frac{1}{m(I)} \int_I f(y) - f(x)\ dy < \frac{1}{m(I)} \int_I \varepsilon \to 0.$$

32.2 Lebesgue Differentiation and Density Theorems

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

$$\lim_{\substack{m(B) \to 0 \\ x \in B}} \int \frac{1}{m(B)} \int_B f(y)\ dy = f(x) \text{ a.e.}$$

Note: although it's not obvious at first glance, this really is a theorem about differentiation.

Corollary (Lebesgue Density Theorem): For any measurable set $E \subseteq \mathbb{R}^n$, we have

$$\lim_{r \to 0} \frac{m(E \bigcap B_r(x))}{m(B_r(x))} = 1 \ a.e.$$

Proof: Let $f = \chi_E$ in the theorem.

We want to show

$$Df(x) \coloneqq \limsup_{\substack{m(B)\to 0 \\ x\in B}} \left| \frac{1}{m(B)} \int_B (f(y) - f(x))\, dy \right| \to 0$$

Note that we can replace $\limsup \cdots$ with

$$\lim_{\varepsilon\to 0} \sup_{\substack{0\leq m(B)\leq\varepsilon \\ x\in B}} \cdots,$$

which is well defined as it is a decreasing sequence of numbers bounded below by zero.

Next we'll introduce that *Hardy-Littlewood Maximal Function*, given by

$$Mf(x) \coloneqq \sup_{x\in B} \frac{1}{m(B)} \int_B |f(y)|\, dy$$

> *Exercise*: show that Mf is a measurable function. (Hint: it's easy to show that the appropriate preimage is open.)

Theorem (Hardy-Littlewood Maximal Function Theorem): Let $f \in L^1(\mathbb{R}^n)$, then

$$m(x \in \mathbb{R}^n \mid Mf(x) > \alpha) \leq \frac{3^n}{\alpha} \|f\|_1.$$

> Idea: if you look at all balls intersecting a given ball of radius α, the worst case is that the other ball doesn't intersect and is of the same radius. But then you can draw a ball of radius 3α and cover every such intersecting ball.

> Exercise: As a corollary, $Mf(x) < \infty$ *a.e.*

This is called a *weak type* estimate, compared to a strong type $\|Mf\|_1 \leq C\|f\|_1$. Note that by Chebyshev, a strong estimate would imply the weak one because

$$m(\left\{x \mid mf(x) > \alpha\right\}) \leq \frac{1}{\alpha} \|Mf\|_1 \not\leq \frac{C}{\alpha} \|f\|_1,$$

which is an inequality that doesn't hold (hence the theorem) because there is an L^1 function for which Mf is *not* L^1.

Proof of differentiation theorem: The goal is to show $Df(x) = 0$ a.e.

We will show that $m(\left\{x \mid Df(x) > \alpha\right\}) = 0$ for all $\alpha > 0$.

Some facts:

1. If g is continuous, then $Dg(x) = 0$ a.e. by uniform convergence.
2. $$D(f_1 + f_2)(x) \leq Df_1(x) + Df_2(x)$$

   by applying the triangle inequality and distributing the $\limsup$.

3.

$$Df(x) \leq Mf(x) + |f(x)|$$

Fix an α and fix an ε. Choose a continuous g such that $\|f-g\|_1 < \varepsilon$.

Writing $f = f - g + g$, we have

$$\begin{aligned} Df(x) &\leq D(f-g)(x) + Dg(x) \\ &= D(f-g)(x) + 0 \\ &\leq M(f-g)(x) + |(f-g)(x)|., \end{aligned}$$

Then

$$Df(x) \geq \alpha \implies M(f-g)(x) \geq \frac{\alpha}{2}$$

or

$$|(f-g)(x)| \geq \frac{\alpha}{2}.$$

So we have

$$\left\{x \;\middle|\; Df(x) > \alpha\right\} \subseteq \left\{x \;\middle|\; M(f-g)(x) > \frac{\alpha}{2}\right\} \bigcup \left\{x \;\middle|\; |f(x) - g(x)| > \frac{\alpha}{2}\right\}.$$

Applying measures turns this into an inequality.

But then applying the maximal function theorem, we have

$$\begin{aligned} m(\left\{x \;\middle|\; Df(x) > \alpha\right\}) &\leq \frac{3^n}{\alpha/2}\|f-g\|_1 + \frac{2}{\alpha}\|f-g\|_1 \\ &\leq \varepsilon\left(\frac{2(3^n+1)}{\alpha}\right). \end{aligned}$$

■

> Note that somehow proving the maximal function theorem here really paved the way, and allows some generalization. Here we computed an average over a solid ball, but there is a notion of surface measure, so we can consider averaging over the surface of spheres, which can include more exotic objects like spheres in $\mathbb{Z}^d$.

Proof of HL Maximal Function Theorem: Let

$$E_\alpha := \left\{x \;\middle|\; Mf(x) > \alpha\right\}.$$

If $x \in E_\alpha$, then it follows that there is a B_x such that

$$\frac{1}{m(B_x)} \int_{B_x} |f(y)|\, dy > \alpha \iff m(B_x) < \frac{1}{\alpha}\int_{B_x} |f(y)|\, dy.$$

Note that if E_α were compact, there would only be finitely many such balls, so let $K \subseteq E_\alpha$ be a compact subset. We will be done if we can show that

$$m(K) < \frac{3^n}{\alpha} \|f\|_1,$$

since we can always find a compact K such that $m(E_\alpha \setminus K)$ is small.

There exists a finite collection $\{B_k\}^N$ such that each $B_k = B_x$ for some $x \in E_\alpha$, $K \subseteq \bigcup B_k$, and

$$m(B_k) \leq \frac{1}{\alpha} \int_{B_k} |f(y)| \, dy.$$

Supposing that the B_k were disjoint (which they are not!), then we would be done since

$$\begin{aligned} m(K) &\leq \sum m(B_k) \\ &\leq \frac{1}{\alpha} \sum \int_{B_k} |f(y)| \, dy \\ &\leq \frac{1}{\alpha} \int_{\mathbb{R}^n} |f(y)|. \end{aligned}$$

Lemma (The Vitali Covering Lemma): Given any collection of balls $B_1, \cdots, B_N$, there exists a sub-collection $A_1, \cdots, A_M$ which are disjoint with

$$m(\bigcup^N B_k) \leq 3^n \sum^M m(A_j).$$

> Note that this follows directly from picking the largest ball first, then picking further balls that avoid everything already picked and are chosen in decreasing order of size. The 3^n factor comes from the earlier fact that tripling the radius covers everything you didn't pick.

But now we can replace B_k with such a sub-collection A_k in the above set of inequalities, which proves the theorem. $\blacksquare$.

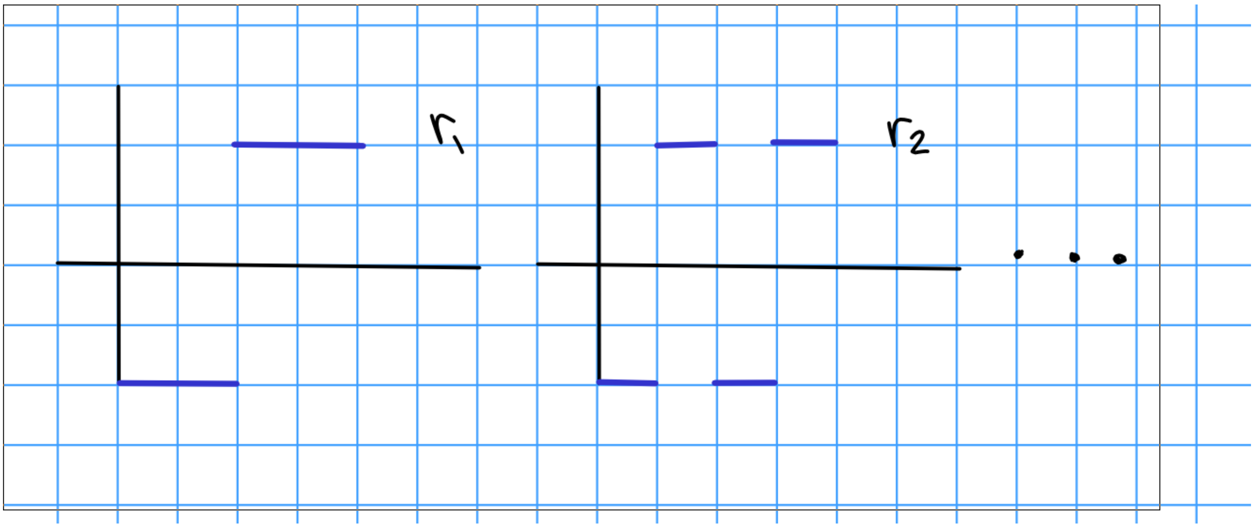
33 Tuesday December 3rd

Lebesgue Differentiation Theorem: If $f \in L_{loc}(\mathbb{R}^n)$, then

$$\lim_{m(B) \to 0; x \in B} \frac{1}{m(B)} \int_B f(y) \, dy = f(x) \quad \text{for almost every } x$$

33.1 Rademacher Functions

Definition: The *Rademacher functions* are given by functions $r_n : [0,1] \to \{-1, 1\}$:



Note the self-similar nature: r_2 just does r_1 on each half-interval. Now lift these to $\mathbb{R}^n$ by taking products.

Facts:

$\{r_n\}$ forms an orthonormal system in $L^2([0,1])$, and

$$\int_0^1 r_n(x)\ dx = 0$$
$$\int_0^1 r_n(x) r_m(x)\ dx = 0 \quad \text{if } n \neq m$$
$$\int_0^1 r_n(x)^2\ dx = 1 \quad \forall n.$$

Think of these as modeling a random process, like coin flips.

Consider

$$a_n(t) := \frac{1}{2}(r_n(t) + 1),$$

which essentially sends $-1 \to 0$ and $1 \to 1$ in each r_n.

Note that when $t = \frac{1}{4}$, we have $\{a_n(t)\} = [0, 0, 1, 1, 1, \cdots]$, which is the binary expansion of t.

If we define $S_N(t) = \sum_{n=1}^{N} r_n(t)$, we have $\int_0^1 S_N(t) = 0$ and $\int_0^1 S_N(t)^2 = n$, which says that the expected value is zero (as many heads as tails) and the variance is additive.

33.2 Law of Large Numbers

Strong Law of Large Numbers:

$$\frac{S_N(t)}{N} \to 0 \text{ as } N \to \infty \text{ for almost every } t \in [0,1]$$

There is in fact a stronger version:

$$\forall \varepsilon > 0, \quad \frac{S_N(t)}{N^{\frac{1}{2}-\varepsilon}} \to 0 \text{ as } N \to \infty \text{ for almost every } t \in [0,1]$$

This is a consequence of the following theorem:

Theorem:

$$\sum |a_n|^2 < \infty \implies \sum a_n r_n(t) < \infty \text{ for almost every } t.$$

Think about $a_n = \dfrac{1}{n}$ or $a_n = \dfrac{1}{n^{\frac{1}{2}+\varepsilon}}$.

Proof of why this theorem implies the Strong Law of Large Numbers:

Let $\dfrac{1}{2} < \gamma < \dfrac{1}{2} + \varepsilon$, and let $a_n = \dfrac{1}{n^\gamma}$ and $b_n = n^\gamma$. Let $\tilde{S}_N = \sum^N a_n r_n$ and

$$\begin{aligned}
S_N &= \sum^N r_n \\
&= \sum^N a_n r_n b_n \\
&= \tilde{S}_N b_{n+1} + \sum_{n=1}^N \tilde{S}_N (b_n - b_{n+1}) \quad \text{by summation by parts} \\
&\leq_{abs} O(1) O(N^\gamma) B_N = O(N^\gamma) + O(N^\gamma),
\end{aligned}$$

where we use the fact that $b_{n+1} > b_n$ just makes the absolute value switch signs, $\tilde{S}_N$ is bounded by a constant, and the resulting sum telescopes. ■

Proof of theorem:

We want to show $\tilde{S}_N = \sum a_n r_n \to f$ in L^2 since the $\{r_n\}$ form an orthonormal system. We have

$$\begin{aligned}
\left\| \tilde{S}_N - \tilde{S}_M \right\|^2 &= \left\| \sum_{n=M+1}^N a_n r_n \right\|^2 \\
&= \sum \|a_n r_n\|^2 \quad \text{by Pythagoras} \\
&= \sum |a_n|^2.
\end{aligned}$$

Claim: $\tilde{S}_N \to f$ almost everywhere.

Proof of claim: Consider $\mathbb{E}_N(f)(t) := \dfrac{1}{m(I)} \displaystyle\int_I f(y)\, dy$ where I is an interval of length 2^{-N} containing t. This replaces f with its average on this interval, and turns out to equal the conditional expectation.

Exercise: $\mathbb{E}_N(f) = \tilde{S}_N$. *Hint: use the fact that*

$$\mathbb{E}_N(r_n) = \begin{cases} r_n & n > N \\ 0 & n \leq N. \end{cases}$$

But then by the Lebesgue differentiation theorem, we have $\mathbb{E}_N(f)(t) \to f(t)$ almost everywhere.

So then $\mathbb{E}_N(\tilde{S}_n) = \tilde{S}_n$ if $n > N$, so just use the fact that $\tilde{S}_N \to f$ in L^2 to complete the argument.

■

That's the end of the course!

34 Appendix

An alternative characterization of **uniform continuity**:

$$\left\|\tau_y f - f\right\|_u \to 0 \text{ as } y \to 0$$

Lemma: Measurability is not preserved by homeomorphisms.

Counterexample: there is a homeomorphism that takes that Cantor set (measure zero) to a fat Cantor set

34.1 Undergraduate Analysis Review

- Some inclusions on the real line:

  Differentiable with a bounded derivative $\subset$ Lipschitz continuous $\subset$ absolutely continuous $\subset$ uniformly continuous $\subset$ continuous
  Proofs: Mean Value Theorem, Triangle inequality, Definition of absolute continuity specialized to one interval, Definition of uniform continuity

- **Bolzano-Weierstrass**: Every bounded sequence has a convergent subsequence.

- **Heine-Borel**:

$$X \subseteq \mathbb{R}^n \text{ is compact} \iff X \text{ is closed and bounded.}$$

- **Baire Category Theorem:** If X is a complete metric space, then

- For any sequence $\{U_k\}$ of open, dense sets, $\bigcap_k U_k$ is also dense.

- X is *not* a countable union of nowhere-dense sets

- **Nested Interval Characterization of Completeness:** $\mathbb{R}$ being complete $\implies$ for any sequence of intervals $\{I_n\}$ such that $I_{n+1} \subseteq I_n$, $\bigcap I_n \neq \emptyset$.

- **Convergence Characterization of Completeness:** $\mathbb{R}$ being complete is equivalent to "absolutely convergent implies convergent" for sums of real numbers.

- Compacts subsets $K \subseteq \mathbb{R}^n$ are also *sequentially compact*, i.e. every sequence in K has a convergent subsequence.

- **Urysohn's Lemma:** For any two sets A, B in a metric space or compact Hausdorff space X, there is a function $f : X \to I$ such that $f(A) = 0$ and $f(B) = 1$.
- Continuous compactly supported functions are
  - Bounded almost everywhere
  - Uniformly bounded
  - Uniformly continuous

  Proof:

  3.3. **Theorem.** A continuous function on a compact set is uniformly continuous.

  Proof. Assume D compact and $f : D \to \mathbb{R}$ continuous. Given $\epsilon > 0$ we need to find $\delta(\epsilon) > 0$ such that if $x, y \in D$ and $|x - y| < \delta(\epsilon)$, then $|f(x) - f(y)| < \epsilon$.
  From the definition of continuity, given such $\epsilon > 0$ and $x \in D$, there exists $\delta_x(\epsilon)$ such that if $|y - x| < \delta_x(\epsilon)$, then $|f(y) - f(x)| < \epsilon$. Clearly $D \subseteq \cup_{x \in D} B_x(\frac{1}{2}\,\delta(\epsilon/2))$. From this open covering we can extract a finite subcovering (D is compact!), meaning there must exists finitely many $x_1, x_2, \ldots, x_N \in D$ such that $D \subseteq \cup_{i=1}^{N} B_{x_i}(\frac{1}{2}\,\delta_{x_i}(\epsilon/2))$.
  Let now $\delta(\epsilon) = \min\{\frac{1}{2}\,\delta_{x_1}(\epsilon/2), \ldots, \frac{1}{2}\,\delta_{x_N}(\epsilon/2)\}$. We will show that $\delta(\epsilon)$ does the job.
  Take $y, z \in D$ arbitrary such that $|y - z| < \delta(\epsilon)$. The idea is that y will be near some x_j, which in turn places z near that same x_j. But that forces both $f(y)$, $f(z)$ to be close to $f(x_j)$ (by continuity at x_j), and hence close to each other.
  Since $y \in D$, there must exist some j, $1 \leq j \leq N$ such that $y \in B_{x_j}(\frac{1}{2}\delta_{x_j}(\epsilon/2))$. Thus

  - $|y - x_j| < \frac{1}{2}\,\delta_{x_j}(\epsilon/2)$
  - but $|y - z| < \delta(\epsilon) \leq \frac{1}{2}\,\delta_{x_j}(\epsilon/2)$

  By the triangle inequality it follows that $|z - x| < \delta_{x_j}(\epsilon/2)$. So y, z are within $\delta_{x_j}(\epsilon/2)$ of x. This implies that

  - $|f(y) - f(x_j)| < \epsilon/2$
  - $|f(z) - f(x_j)| < \epsilon/2$

  By the triangle inequality once again we have $|f(y) - f(z)| < \epsilon$.

- Uniform convergence allows commuting sums with integrals
- Closed subsets of compact sets are compact.
- Every compact subset of a Hausdorff space is closed
- Showing that a series converges: *(Todo)*

34.2 Big Counterexamples

34.2.1 For Limits

- Differentiability $\implies$ continuity but not the converse:
  - The Weierstrass function is continuous but nowhere differentiable.
- f continuous does not imply f' is continuous: $f(x) = x^2 \sin(x)$.
- Limit of derivatives may not equal derivative of limit:

$$f(x) = \frac{\sin(nx)}{n^c} \text{ where } 0 < c < 1.$$

  - Also shows that a sum of differentiable functions may not be differentiable.

- Limit of integrals may not equal integral of limit:

$$\sum \mathbb{1}\left[x = q_n \in \mathbb{Q}\right].$$

- A sequence of continuous functions converging to a discontinuous function:

$$f(x) = x^n \text{ on } [0,1].$$

- The Thomae function *(todo)*

34.2.2 For Convergence

- Notions of convergence:
  1. Uniform
  2. Pointwise
  3. Almost everywhere
  4. In norm

Uniform $\implies$ pointwise $\implies$ almost everywhere.

See Section 17.3.

Almost everywhere convergence does not imply L^p convergence for any $1 \leq p \leq \infty$

See notes section 1

Sequences $f_k \stackrel{a.e.}{\to} f$ but $f_k \stackrel{L^p}{\not\to} f$:

- For $1 \leq p < \infty$: The skateboard to infinity, $f_k = \chi_{[k,k+1]}$.

  Then $f_k \stackrel{a.e.}{\to} 0$ but $\|f_k\|_p = 1$ for all k.

  Converges pointwise and a.e., but not uniformly and not in norm.

- For $p = \infty$: The sliding boxes $f_k = k \cdot \chi_{[0,\frac{1}{k}]}$.

  Then similarly $f_k \stackrel{a.e.}{\to} 0$, but $\|f_k\|_p = 1$ and $\|f_k\|_\infty = k \to \infty$

  Converges a.e., but not uniformly, not pointwise, and not in norm.

The Converse to the DCT does not hold

L^p boundedness does not imply a.e. boundedness.

I.e. it is not true that $\lim \int f_k = \int f$ implies that $\exists g \in L^p$ such that $f_k < g$ a.e. for every k.

Take

- $b_k = \displaystyle\sum_{j=1}^{k} \frac{1}{j} \to \infty$
- $f_k = \chi_{[b_k, b_{k+1}]}$

Then

- $f_k \overset{a.e.}{\to} f = 0$,
- $\displaystyle\int f_k = \frac{1}{k} \to 0 \implies \|f_k\|_p \to 0$,
- $\displaystyle 0 = \int f = \lim \int f_k = 0$
- But $g > f_k \implies g > \|f_k\|_\infty = 1$ a.e. $\implies g \notin L^p(\mathbb{R})$.

34.3 Errata

- **Equicontinuity**: If $\mathcal{F} \subset C(X)$ is a family of continuous functions on X, then $\mathcal{F}$ *equicontinuous* at x iff

$$\forall \varepsilon > 0 \; \exists U \ni x \text{ such that } y \in U \implies |f(y) - f(x)| < \varepsilon \quad \forall f \in \mathcal{F}.$$

- **Arzela - Ascoli 1**: If $\mathcal{F}$ is pointwise bounded and equicontinuous, then $\mathcal{F}$ is totally bounded in the uniform metric and its closure $\overline{\mathcal{F}} \in C(X)$ in the space of continuous functions is compact.
- **Arzela - Ascoli 2**: If $\{f_k\}$ is pointwise bounded and equicontinuous, then there exists a continuous f such that $f_k \overset{u}{\to} f$ on every compact set.

Example: Using Fatou to compute the limit of a sequence of integrals:

$$\lim_{n\to\infty} \int_0^\infty \frac{n^2}{1+n^2x^2} e^{-\frac{x^2}{n^3}} \, dx \overset{\text{Fatou}}{\geq} \int_0^\infty \lim_{n\to\infty} \frac{n^2}{1+n^2x^2} e^{-\frac{x^2}{n^3}} \, dx \to \int \infty.$$

Note that MCT might work, but showing that this is non-decreasing in n is difficult.

Lemma:

$$f_k \overset{a.e.}{\to} f, \quad \|f_k\|_p \leq M \implies f \in L^p \text{ and } \|f\|_p \leq M.$$

Proof: Apply Fatou to $|f|^p$:

$$\int |f|^p = \int \liminf |f_k|^p \leq \liminf \int |f_k|^p = M.$$

Lemma: If f is uniformly continuous, then

$$\|\tau_h f - f\|_p \overset{L^p}{\to} 0 \quad \text{for all } p.$$

Lemma: $\|\tau_h f - f\|_p \to 0$ for every p.

- i.e. "Continuity in L^1" holds for all L^p.
- i.e. Translation operators are continuous.

Proof: Take $g_k \in C_c^0 \to f$, then g is uniformly continuous, so

$$\|\tau_h f - f\|_p \le \|\tau_h f - \tau_h g\|_p + \|\tau_h g - g\|_p + \|g - f\|_p \to 0.$$

Lemma: For $f \in L^p, g \in L^q$, $f * g$ is uniformly continuous.

Proof: Use Young's inequality

$$\|\tau_h(f*g) - f*g\|_\infty = \|(\tau_h f - f)*g\|_\infty \le \|\tau_h f - f\|_p \|g\|_q \to 0.$$

Lemma: If $\displaystyle\int f\phi = 0$ for every $\phi \in C_c^0$, then $f = 0$ almost everywhere.

Proof: Let A be an interval, choose $\phi_k \to \chi_A$, then $\displaystyle\int f\chi_A = 0$ for all intervals. So this holds for any Borel set A. Then just take $A_1 = \{f > 0\}$ and $A_2 = \{f < 0\}$, then $\displaystyle\int_{\mathbb{R}} f = \int_{A_1} f + \int_{A_2} f = 0$.

34.4 The Fourier Transform

Some Useful Properties:

$$\begin{aligned}
\widehat{f*g}(\xi) &= \hat{f}(\xi)\cdot\hat{g}(\xi) \\
\widehat{\tau_h f}(\xi) &= e^{2\pi i\xi\cdot h}\widehat{f}(\xi) \\
\widehat{e^{2\pi i\xi\cdot h} f}(\xi) &= \tau_{-h}\widehat{f}(\xi) \\
\widehat{f\circ T}(\xi) &= |\det T|^{-1}(\hat{f}\circ T^{-t})(\xi) \\
\frac{\partial}{\partial\xi}\widehat{f}(\xi) &= -2\pi i\cdot\widehat{\xi f}(\xi) \\
\widehat{\frac{\partial}{\partial\xi} f}(\xi) &= 2\pi i\xi\cdot\widehat{f}(\xi).
\end{aligned}$$

Some Useful Transform Pairs:

Dirichlet:	$\chi_{\{-\frac{1}{2}\le x\le\frac{1}{2}\}}$	$\Longleftrightarrow$	$\operatorname{sinc}(\xi)$
Fejer:	$\chi_{\{-1\le x\le 1\}}(1-\|x\|)$	$\Longleftrightarrow$	$\operatorname{sinc}^2(\xi)$
Poisson:	$\dfrac{1}{\pi}\dfrac{1}{1+x^2}$	$\Longleftrightarrow$	$e^{2\pi\|\xi\|}$
Gauss-Weierstrass:	$e^{-\pi x^2}$	$\Longleftrightarrow$	$e^{-\pi\xi^2}$.