Real Analysis

D. Zack Garza

November 12, 2019

Contents

1 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 and 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

1.1 Big Theorems

- **Bolzano-Weierstrass**: every bounded sequence has a convergent subsequence.
- **Heine-Borel**: in $\mathbb{R}^n$, X is compact $\iff$ X is closed and bounded.
- **Monotone Convergence Theorem (MCT)**: If $f_n \in L^+$ and $f_n \nearrow f$ a.e., then $\lim \int f_n = \int \lim f_n = \int f$.
- **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 $\lim \int f_n = \int \lim f_n = \int f$.
- **Fatou**: If $f_n \in L^1$ (with no other real conditions), then $\int \liminf f_n \leq \liminf \int f_n$.

1.2 Big Counterexamples

- 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: $f(x) = \frac{\sin(nx)}{n^c}$ 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: $f(x) = x^n$ on $[0, 1]$.

2 Lecture 1 (Thu 15 Aug 2019 11:04)

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.

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

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

DefinitionA 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:

DefinitionA 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 countable union above. For 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 something is measure zero exactly when it can be covered by intervals whose lengths sum to less than ε.

DefinitionDefinition: 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

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

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.

2.2 $\mathbb{R}$ is not small

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

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

Theorem: $\mathbb{R}$ 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 $a < b$. This 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.

Proof (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. $\square$

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.

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 one. Equivalently, the rationals are not a G_δ set.

Some facts:

- The pointwise limit of continuous functions has a meager set of discontinuities.
- If f is integrable, the set of discontinuities is null.
- If f is monotone, they are 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, the discontinuities are null.

Theorem: Let $I = [a, b]$. If $I \subseteq \bigcup_{i=1}^{\infty} I_i$, then $|I| \leq \sum_{i=1}^{\infty} |I_i|$.

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 the reals are uncountable!

3 Lecture 2 (Thursday)

(Find notes for first 15 mins)

- 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 $\{x_n\} \to x$.

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

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

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.

Exercise: Prove theorems 1 and 2.

3.1 Proof of Theorem 3

Proof of theorem 3: 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 $F_n = \{x \ni \omega_f(x) \geq 0\}$ 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

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

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 $|f_i(x) - f_j(x)| \le 1/5n$ for all $i, j \ge k$. So let $i \to \infty$, so $|f(x) - f_j(x)| \le 1/5n$ for all $j \ge k$.

Now for any $x \in J$, there exists some interval $I(x) \subseteq J$ depending on x such that $|f(y) - f_k(x)| \le 2/5n$. (where we will rewrite this as $f(y) - f_k(y) + f_k(y) - f_k(x)$).

This implies that $\omega_f(x) \le 4/5n$. $\square$.

3.2 Proof of Theorem 4

Suppose that $f : [a, b] \to \mathbb{R}$ is bounded. Recall that f is Riemann integrable iff for any ε there exists a partition $P_\varepsilon = \{a = x_1 \le x_2 \le \cdots x_n = b\}$ of $[a, b]$ such that $U(f, P_\varepsilon) - L(f, \varepsilon) \le \varepsilon$, where this expression is equal to

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

($\Rightarrow$): 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 $1/n$,

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

($\Leftarrow$): 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 $\{I_n\}^N$ with $\sum |I_n| \le \varepsilon$.

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.

4 Lecture 3 (Tuesday)

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 if 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. Suppose it was, so $\mathbb{R} \setminus \mathbb{Q} = \bigcup_{n \in \mathbb{N}} F_n$ with A_n closed. Then $\mathbb{R} = (\bigcup_n F_n) \bigcup (\bigcup_i \{q_i\})$ 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: Showing that F_n are nowhere dense by constructing a sequence of elements in F_n that converges to an element in $F_n^c \subset \mathbb{Q}$.

4.1 Riemann Integration

Some good 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

$$\mathcal{L} = \{f \ni f \text{ is Lebesgue integrable }\},$$

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 $\int f_n \to \int f$.

4.2 Measure Theory

4.2.1 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 $m(\coprod_{i\in\mathbb{N}} E_i) = \sum_{i\in\mathbb{N}} m(E_i)$.
- 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. There exists a non-measurable set: 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 $\mathbb{Q} \bigcap [-1, 1] = \{q_j\}$, 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 $[0, 1) \subseteq \bigcup_j N_j \subseteq [-1, 2]$. But then by taking measures, $1 \leq \sum_j m(N_j) \leq 3$. Using $m(N_i) = m(N)$.

But then $m(N) = 0 \implies 1 > m(N)$, and if $m(N) = \varepsilon$ then $\sum m(N) = \sum \varepsilon > 3$.

Any open set in $\mathbb{R}$ can be written as a *countable* union of intervals (Exercise). But how do you go to closed sets, or to $\mathbb{R}^n$? Any open set in $\mathbb{R}^n$ can be written as an *almost disjoint* union of closed cubes. We can then attempt to ascribe measure by approximating an open set from the inside by cubes, although it's not clear that this is unique (although it is).

5 Lecture 4 (Thursday)

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.

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:

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

i.e. 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

Example: 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 $m_*(Q) = |Q|$.

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

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! It is not the case that $m_*(G) - m_*(E) \neq m(G \setminus E)$. 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. If $E_1, E_2 \subseteq \mathbb{R}^n$ and

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

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

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.

Proof:

1. Obvious since any covering of E_2 is also a covering of E_1. We're taking infimums over *large* collections of sets, so it could only get smaller.

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

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

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 $G = \bigcup S_j$, which is open. Now using subadditivity, we have

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

4. Just need to show $m_*(E_1 \bigcup E_2) \leq m_*(E_1) + m_*(E_2)$, since the reverse direction follows from (2). 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 $Y = \left\{ n + \frac{1}{2n} \ni n \in \mathbb{N} \right\}$.

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

6 Lecture 5 (Wed 28 Aug 2019, 1:30 PM)

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

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

where the inf 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 $0 < \delta < \text{dist}(E_1, E_2)$ 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 $\text{diam}(Q_i) < \delta$. Then each Q_i can intersect at most one of E_1, E_2. Let $J_1 = \{j \ni Q_j \bigcap E_1 \neq \emptyset\}$ and $J_2 = \{j \ni Q_j \bigcap E_2 \neq \emptyset\}$. Note that J_1, J_2 are disjoint, and we have $E_1 \subseteq \bigcup_{j \in J_1} Q_j$ and likewise for E_2.

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.

Proof of property (5): Let $\varepsilon > 0$, we will show

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

Shrink 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.

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

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 $\bigcup_{j=1}^{N} \tilde{Q}_j \subseteq E$, 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

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

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

7 Lecture 6

Recall the definition of the Lebesgue measure:

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

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

This satisfies properties (1) through (5).

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

E is said to be *measurable* if

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

Some 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: 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: 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

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

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 $S := \bigcup G_k^c$ 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 $E^c \setminus S \subset G_k \setminus E = G_k \bigcap E^c$ for every k, we have $m_*(E^c \setminus S) \leq m_*(G_k \setminus E) < \frac{1}{k}$ 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 measurable!)

Next time: closed sets are measurable, proof of theorem 2, and characterizations of measurability.

8 Lecture 7

Recall that a set E is *Lebesgue measurable* iff there exists an open set G with $E \subseteq G$ and $m(G \backslash E) < \varepsilon$ 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 $\mathrm{Dist}(F, K) > 0$.

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

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

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 $\mathrm{dist}(K, \bigcup^N Q_j) > 0$. Using (1) and (4),

$$m_*(G) \geq m_*(K \bigcup (\bigcup^N) Q_j) = m(K) = m(\bigcup^N Q_j) = m_*(K) + \sum^N |Q_j|, .$$

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

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

8.1 Characterizations of Measurability

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

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

and so

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

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.

We have the inclusions

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

Qual problem alert!

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

Exercise: show both directions!!

Proof: $\impliedby$ is the easy direction.

$\implies$: 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. $\square$

Recall that if E is measurable, we define the 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 $\bigcup^N K_j \subset E \coloneqq \bigcup^\infty E_j$, so

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

but $\varepsilon > 0$ was arbitrary, so this shows the bounded case.

In general, let

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

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

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

where we've applied the special bounded case in the last two equalities.

9 Lecture n: A Brief Introduction to (Actual) Measure Theory

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_*.

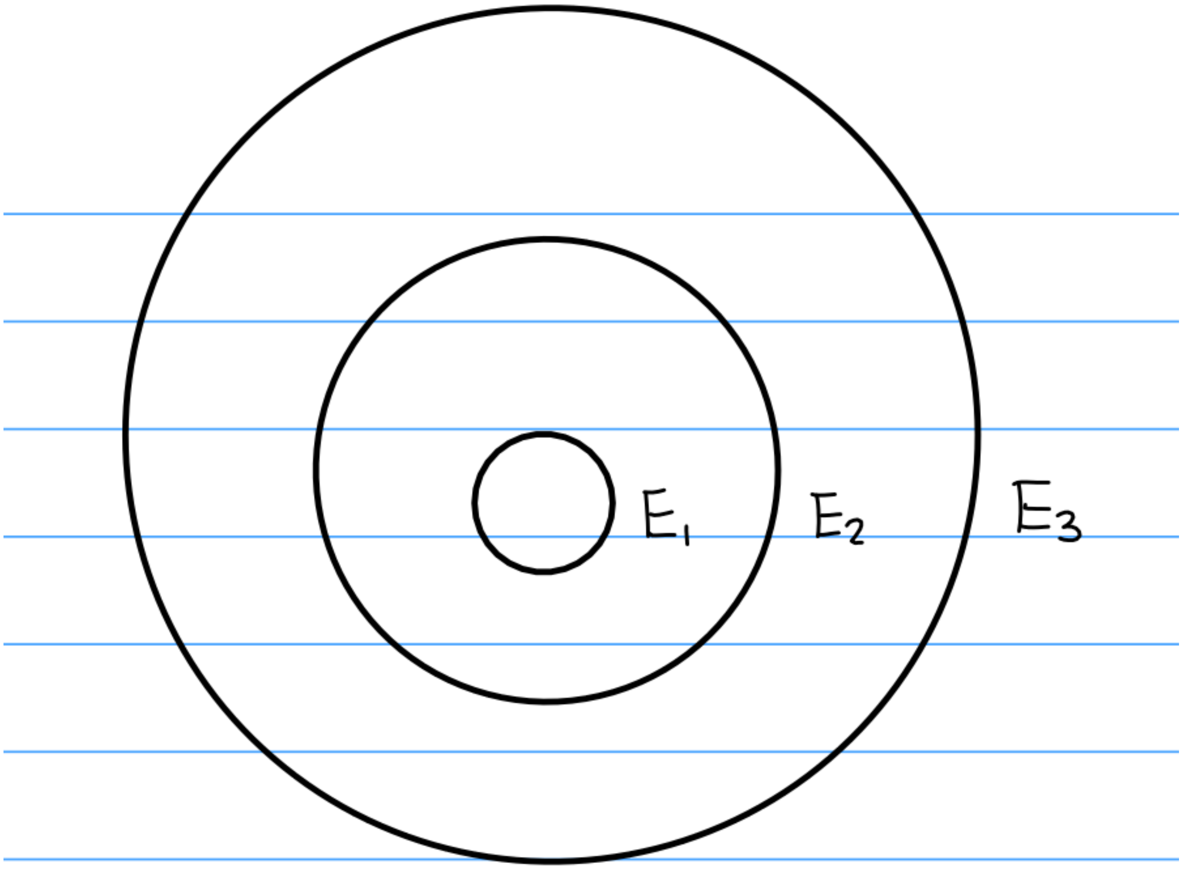
Theorem (Fundamental Theorem of Measure Theory, i.e. continuity from above and below):

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.

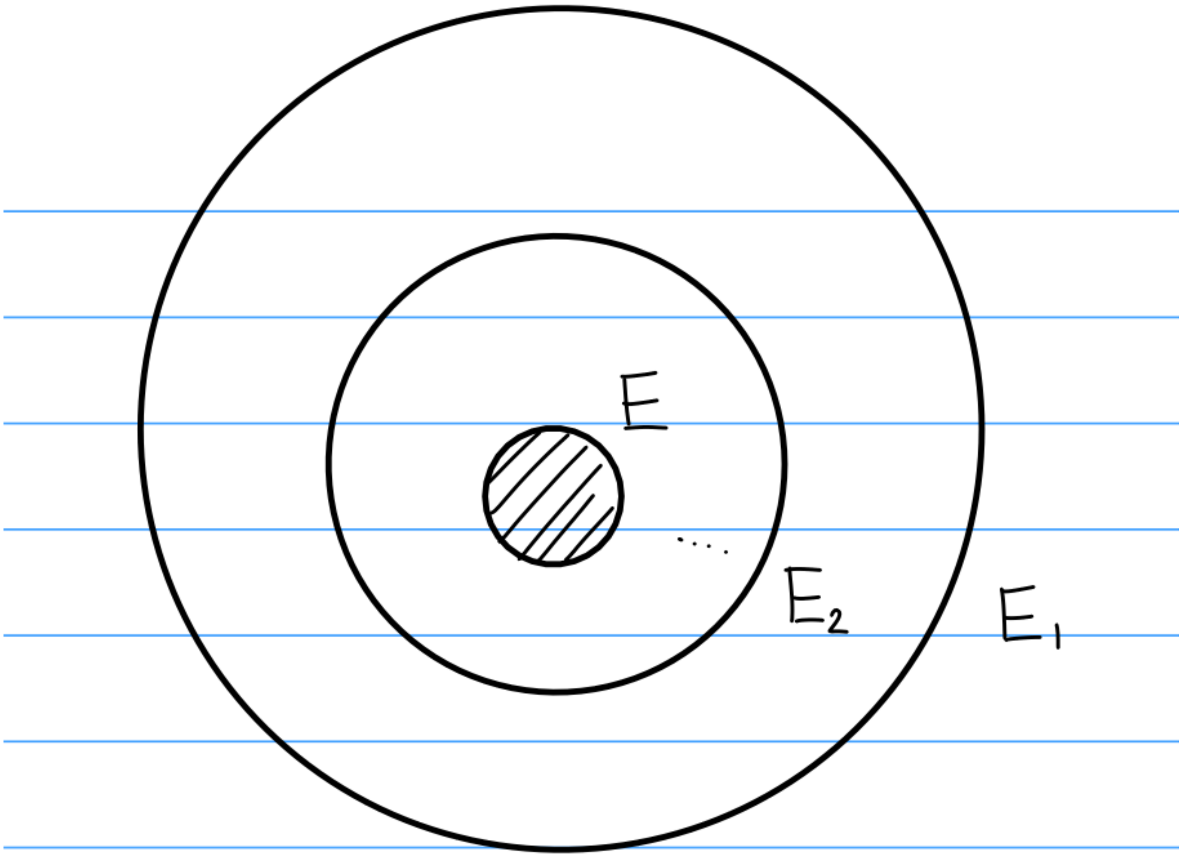
Proof:



(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

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



(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,

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

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

$$\begin{aligned}
& m_*(E) = \inf \{ m_*(G) \ni E \subset G \text{ open} \} \\
\iff & \forall \varepsilon > 0, \ \exists G \supseteq E \ni m_*(G) \leq m_*(E) + \varepsilon.
\end{aligned}$$

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

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

$$\begin{aligned}
& m(E) = \sup \{ m(K) \ni K \subseteq E \text{ and } K \text{ is compact} \} \\
\iff & \forall \varepsilon \exists K \subseteq \ \text{compact} \ \ni m(K) \geq m(E) - \varepsilon.
\end{aligned}$$

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

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

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

Case 2: Suppose E is unbounded. Write $E_j = E \bigcap B(j, 0)$, so $E_j \nearrow E$. Using continuity from below, we have $m(E) = \lim m(E_j)$.

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

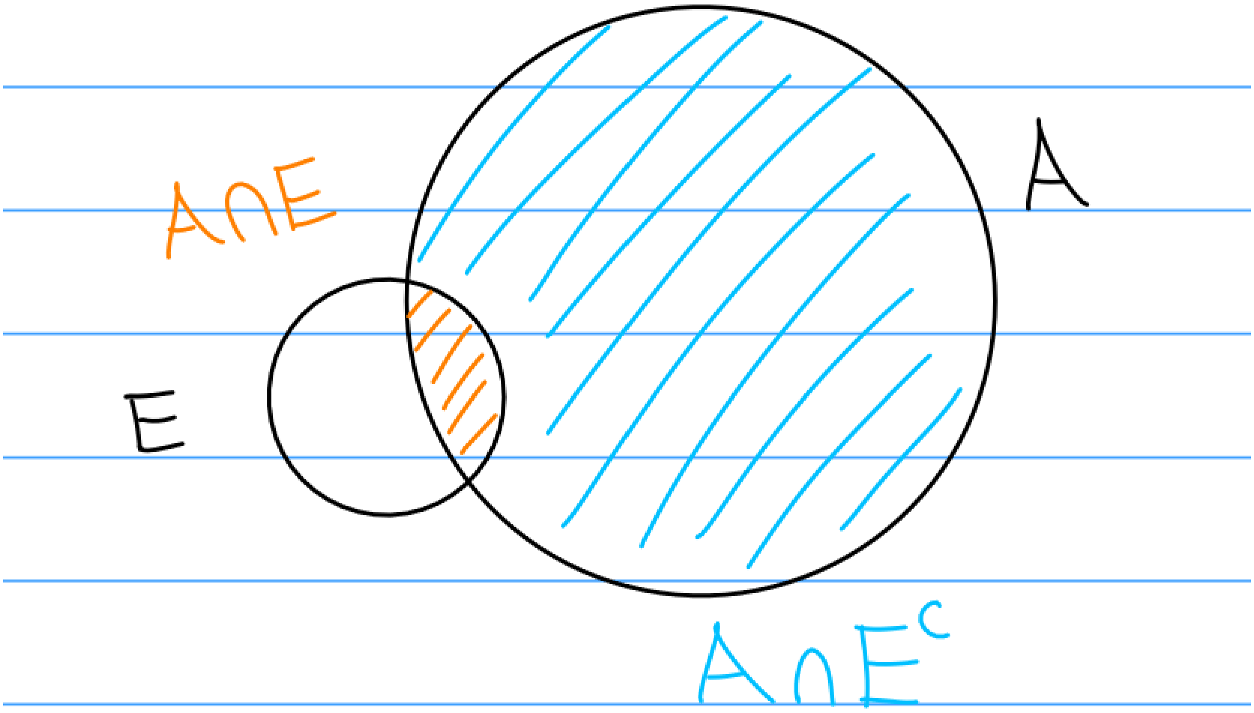


Figure 1: Image

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 $m(K) > M$.

Note: Very similar arguments are often used on the quals.

9.1 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

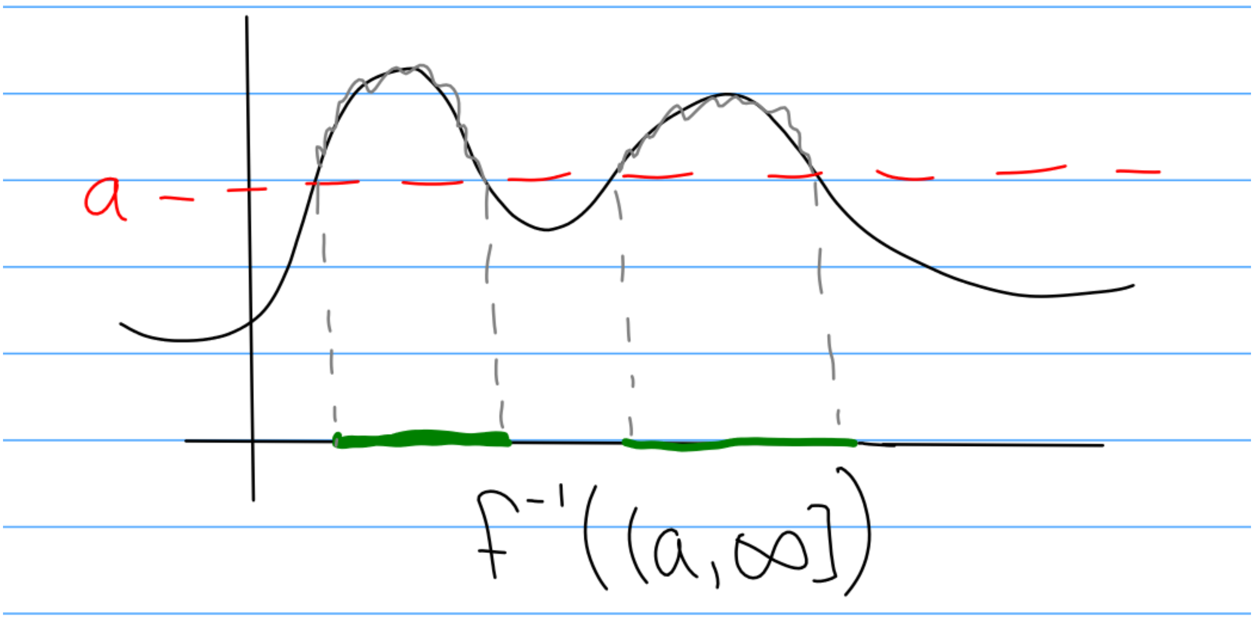


Figure 2: Image

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.

10 Lecture: Measurable Functions

Let $E \subseteq \mathbb{R}^n$ be measurable. Then $f : E \to \mathbb{R} \bigcup \{\pm\infty\}$ is *Lebesgue measurable* iff $\{x \in E \ni f(x) > a\} = f^{-1}((a,\infty])) \in \mathcal{M}$, 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 $\bigcap f^{-1}((a - \frac{1}{k}, \infty])$, 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 is measurable $\iff$ $f^{-1}(G) \in \mathcal{M}$ 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.

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

> Note that this implies that if $f_n \to f$ pointwise and each f_n is measurable, then f is measurable.

Proof of Theorem 2:

Suppose f, g are measurable and finite-valued. Let $a \in \mathbb{R}$ and consider $S = \{x \ni f(x) + g(x) > a\}$. Then $S = \{x \ni f(x) > -g(x) + a\}$, where f is measurable and so is $-g + a$. With the following lemma, we'll be done.

Lemma: If f, h are measurable, then $S = \{x \ni f(x) > h(x)\}$ is always a measurable set. Since $f(x) > h(x)$, there is a rational q such that $h(x) < q < f(x)$. But then $S = \bigcup_{q \in \mathbb{Q}} \{x \ni f(x) > q > h(x)\} = \{f > q\} \bigcap \{h < q\}$, both of which are measurable.

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: $fg = \frac{(f+g)^2-(f-g)^2}{4}$.

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 $S = \{x \ni \sup_n f_n(x) > a\}$ is measurable. Then there is some n for which $f_n(x) > a$, and the claim is that $S = \bigcup_n \{x \ni f_n(x) > a\}$. But this follows formally by just checking set inclusions. So S is a countable union of measurable sets and thus measurable.

Definition: If $f, g : E \to \overline{\mathbb{R}}, \mathbb{C}$ then $f = g$ *almost everywhere* (or $f = g$ a.e.) iff $\{x \ni f(x) \neq g(x)\}$ 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, and thus Z is null.

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" everywhere. A point is that 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*)

11 Lecture qqq

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

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

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.

A *simple function* is a function of the form $s(x) = \sum_{i=1}^N a_j \chi_E(x)$ where E is measurable.

Theorem 1: If $f : E \to [0, \infty]$ is a measurable function (note that this is positive valued), then there exists a sequence of simple functions $\{s_k\}$ such that $s_k(x) \leq s_{k+1}(x)$ for all x and k, and $\lim_{k\to\infty} s_k(x) = f(x)$ for all x.

Corollary: This holds for $f : E \to \overline{\mathbb{R}}$, 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 $\psi_k \to f$ a.e.

Proof: See homework 3, number 1c.

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

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

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 $f(x) := \lim_{k\to\infty}(x)$ exists a.e. and is finite valued. Then the convergence is *almost uniform*, i.e.

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

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 $f_k(x) = k\mathbb{1}\left[0 \le x \le 1 - \frac{1}{k}\right]$. Then $f_k \to \infty \iff 0 \le x < 1,\ 0 \iff x = 1$, but not "almost uniformly".

Theorem (Lusin): If 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 $m(E \setminus F) < \varepsilon$ and $f|_F$ 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.

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

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

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 $|f_k(x) - f(x)| < \frac{1}{j}$ on F_j for all $k \ge k_j$.

So take $F := \bigcap F_j$, which is *closed.* Then by subadditivity, we have $m(E/F) \le \sum m(E \setminus F_j) < \varepsilon$ 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. $\square$

Proof of lemma: Fix ε, α. Define

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

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. $\square$

A note on *convergence in measure*:

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

$$\lim_{k\to\infty} m(\{x \in E \ni |f_k(x) - f(x)| > \alpha\}) = 0. \quad \square$$

How does this relate to pointwise convergence?

Theorem: If $m(E) < \infty$, then $f_k \to f$ *a.e* on $E \iff f_k \to^m f$ 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 a.e.

11.1 Tuesday Lecture

Definition: A simple function $\mathbb{R}^n \to \mathbb{R}$ is a finite $\mathbb{C}$-linear combination of characteristic functions of measurable sets, i.e.

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

Equivalently, $f : \mathbb{R}^n \to \mathbb{C}$ is simple $\iff f$ is measurable and the range of f is a *finite* subset of $\mathbb{C}$.

This lets us write $E_j = f^{-1}(\{a\}_j)$ are all disjoint and the a_j are distinct, and it's not a big jump to also require $\bigcup E_j = \mathbb{R}^n$ by including the set where f is zero. This is *standard representation* of a simple function f.

Theorem (Approximation by Simple Functions): If $f : E \to [0, \infty]$ is measurable, then there exists a sequence of simple functions $\{\phi_i\}_{i=1}^{\infty}$ such that

- $\phi_k(x) \leq \phi_{k+1}(x) \forall x, \forall k$, and
- $\phi_k(x) \to f(x)$ pointwise $\forall x$.

Proof:

First take an initial approximation. For each $k \in \mathbb{N}$, let $Q_k := [-\frac{k}{2}, \frac{k}{2}]^n$. Now let

$$F_k(x) = \begin{cases} f(x) & x \in Q_k \text{ and } f(x) \leq k, \\ k & x \in Q_k \text{ and } f(x) > k, \\ 0 & \text{otherwise.} \end{cases}$$

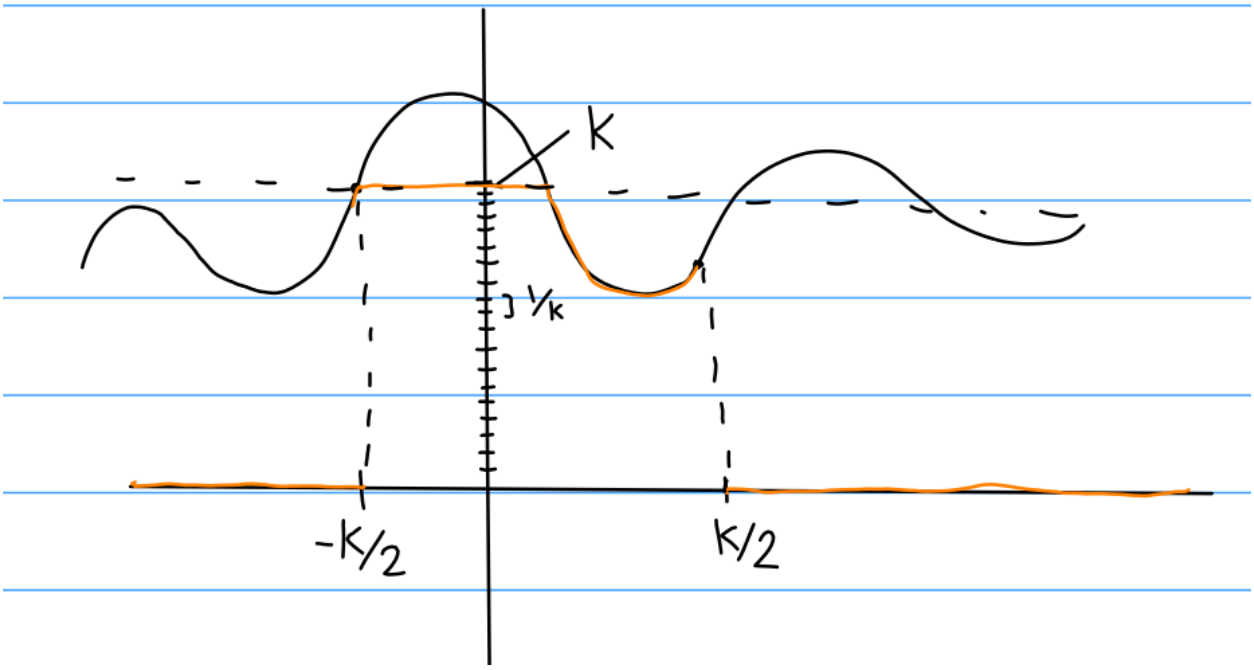


Figure 3: Image

Then $\lim_{k\to\infty} F_k(x) = f(x)$ for all x.

Now break the vertical distance k up into pieces of width $\frac{1}{k}$, and for each $0 \leq j \leq k^2$, define

$$E_{j,k} := \left\{ x \in Q_k \ni \frac{j}{k} < F_k(x) \leq \frac{j+1}{k} \right\}.$$

Now let

$$\phi_k(x) = \sum_{j=0}^{k^2} \frac{j}{k} \chi_{E_{k,j}}.$$

Then $\phi_k(x)$ is a simple function that is increasing, and we have

$$|f(x) - \phi_k(x)| \leq |f(x) - F_k(x)| + |F_k(x) - \phi_k(x)| \leq 0 + \frac{1}{k} \to 0.$$

11.2 Lebesgue Integration

Step 1: Define $\int f$ for f non-negative and simple.

Let

$$L^+ := \{ f : \mathbb{R}^n \to [0, \infty] \ni f \text{ is measurable } \}.$$

If $\phi \in L^+$ is written in the standard representation, i.e. $\phi(x) = \sum_{i=1}^N a_i \chi_{E_i}$, then define

$$\int \phi := \int \phi(x)\ dx = \sum_{j=1}^{N} a_j m(E_j).$$

where we may have $a_j = 0$ and $m(E_j) = \infty$ for some j, so we do this with the understanding that $0 \times \infty = 0$.

Note that if A is measurable, then $\phi\chi_A$ is also simple and

$$\int_A \phi = \int \phi\chi_A.$$

Properties:

1. If $\phi, \psi \in L^+$ are simple, then
    a. If $c \geq 0$ then $\int c\phi$ is defined and equal to $c\int \phi$. (Easy to show)
2. $\int(\psi + \phi) = \int \psi + \int \phi$.
3. If $\psi \leq \phi$, then $\int \psi \leq \int \phi$.
4. The map $A \xrightarrow{\mu_A} \int_A \phi$ is a *measure* (ϕ acts like a density).

Proof of (1): Easy

Proof of (2): Write ϕ, ψ in standard representation, so

$$\begin{aligned}\phi &= \sum a_i \chi_{E_i} \\ \psi &= \sum b_i \chi_{F_i}.\end{aligned}$$

Then note that $E_j = \bigcup_{k=1}^{M} E_i \bigcap F_k$, and each of these terms are disjoint since the F_k were disjoint, and so

$$\begin{aligned}\int \phi &= \sum_{j=1}^{n} a_j m(E_j) = \sum_{j=1}^{n} \sum_{i=1}^{M} a_j m(E_j \bigcap F_k) \\ \int \psi &= \sum_{j=1}^{n} b_j m(F_j) = \sum_{j=1}^{n} \sum_{i=1}^{M} b_j m(E_j \bigcap F_k).\end{aligned}$$

But note that while $\phi = \sum_{j,k} a_j \chi_{E_j \bigcap F_k}$ is not a standard representation, and likewise for ψ, but the integral takes the expected form.

So we have

$$\int \psi + \int \phi = \sum_{j,k} (a_j + b_j) m(E_j \bigcap F_k).$$

But consider the LHS of the original equation. We have $\phi + \psi = \sum_{j,k} (a_j + b_k) \chi_{E_j \bigcap F_k}$, which is almost in standard form. The sets are disjoint, but it may occur that $a_j + b_k$ are not distinct. But if, for example, 2 occurs twice in these values, and there are two sets in the preimage, we can just take their union. Since the sets are disjoint, the measure of the union is the sum of the measures, and so this quantity is still $\int(\psi + \phi)$ as desired.

Proof of (3): If $\phi \leq \psi$, then $a_j \leq b_k$ when $E_j \bigcap F_k \neq \emptyset$. Then,

$$\int \phi = \sum_{j,k} a_i m(E_j \bigcap F_k) \leq \sum_{j,k} b_k m(E_j \bigcap F_k) = \int \psi.$$

Proof of (4): Suppose $A \in \mathcal{M}$, so A is measurable. We want to show $\mu_A := \int_A \phi$ is countably additive.

So suppose $A = \coprod A_k$ is a disjoint union of measurable sets. We want to show $\int_A \phi = \sum_k \int_{A_k} \phi$.

We have

$$\int_A \phi = \sum a_i m(E_i \bigcap A) = \sum_k \sum_j a_j m(E_j \bigcap A_k) = \sum_k \int_{A_k} \phi.$$

where we note that $E_j \bigcap A = \bigcup_k E_j \bigcap A_k$, which is a disjoint union of measurable sets.

Extending the definition to all of L^+:

If $f \in L^+$, we define

$$\int f = \sup \left\{ \int \phi \ni 0 \leq \phi \leq f,\ \phi \text{ simple} \right\}.$$

Note that if f itself is simple, then by property (3), this definition agrees with the previous definition.

We also have

1. $f \leq g \implies \int f \leq \int g$, since we're taking the supremum over a larger set on the RHS.
2. $c \geq 0 \implies \int cf = c \int f$.

Note that it is not obvious that $\int(f+g) = \int f + \int g$ for $f, g \in L^+$.

Incredibly Important Theorem (Monotone Convergence): If $\{f\}_n \subset L^+$ with $f_n \leq f_{n+1}$ for all n and $f_n \to f$ pointwise, then

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

Consequence of this theorem: If $f, g \in L^+$, then $\exists \phi_n \nearrow f, \psi_n \nearrow g$, and so $\phi_n + \psi_n \nearrow f + g$ and

$$\begin{aligned}
\int (f+g) &=_{\text{MCT}} \lim \int (\phi_n + \psi_n) \\
&= \lim (\int \phi_n + \int \psi_n) \qquad \text{since } \phi_n, \psi_n \text{ are simple} \\
&= \lim \int \phi_n + \lim \int \psi_n \\
&=_{\text{MCT}} \int f + \int g..
\end{aligned}$$

12 Thursday Lecture

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 $\int \phi = \sum a_i m(E_i)$ iff this is a standard representation for ϕ.

We then extend to all functions in L^+ by defining $\int f = \sup\{\int \phi \ni 0 \le \phi \le f,\ \phi \text{ simple }\}$.

Some properties:

- $f \le g \implies \int f \le \int g$ (monotonicity, easy to show)
- $\int cf = c \int f$ (also easy to show)
- $\int (f+g) = \int f + \int g$ (less obvious, follows from MCT)

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

Theorem: If $\{f_n\} \subset L^+$, then $\sum \int f_k = \int \sum f_k$. (Exercise, not too tricky.)

Proof of MCT: Given any simple $\phi \in L^+$, 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 $f_k \in L^+ \implies \lim f_k \int L^+$, and $f_k \le f_{k+1} \implies \int f_k \le \int f_{k+1}$, so the limit on the LHS makes sense.

Let $f = \lim f_k$, then $\int f_k \le \int f$ for all k, so $\lim \int f_k \le \int f$. So we need to show that $\lim \int f_k \ge \int f$.

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

Let $E_k = \{x \ni f_k(x) \ge \alpha\phi(x)\}$, 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 $f_k \to f, \alpha\phi < \phi \le f$.

Then

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

But then

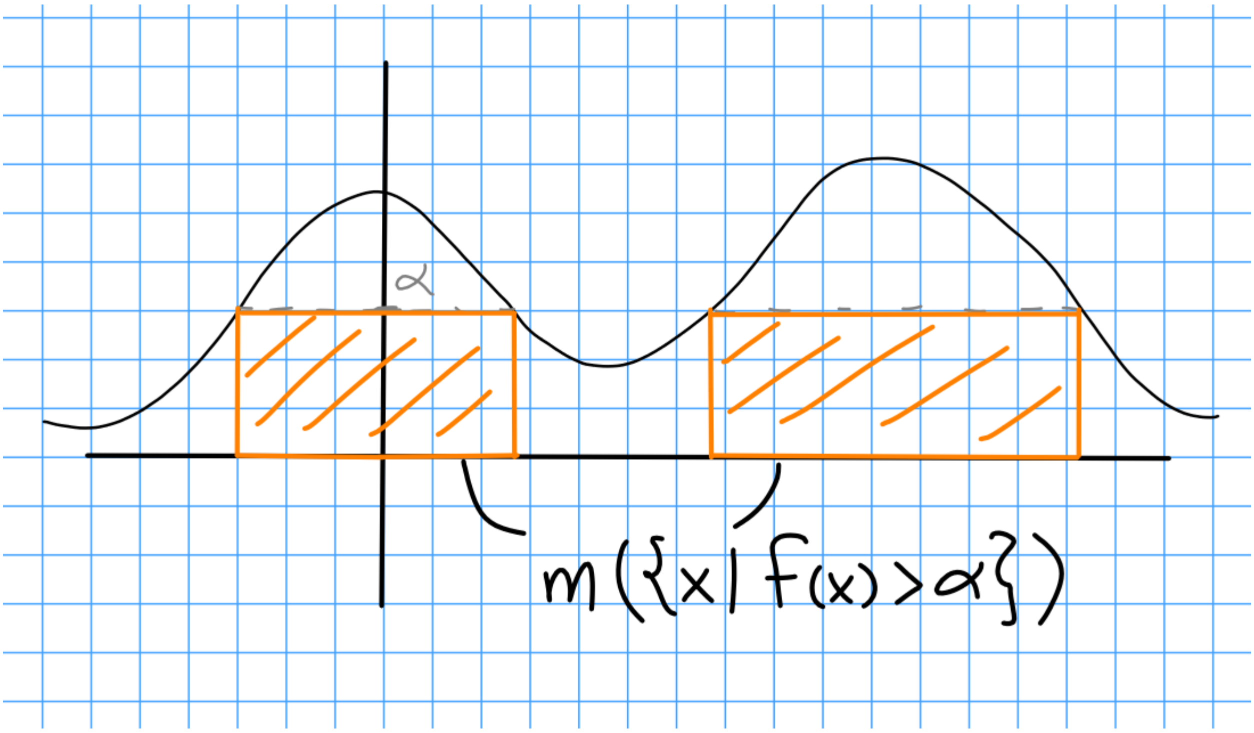
$$\lim \int f_k \ge \lim \alpha \mu_\phi(E_k) = \alpha \mu_\phi(\mathbb{R}^n) = \alpha \int \phi \quad \forall \alpha, \forall \phi.$$

So $\lim \int f_k \ge \int \phi$ for all ϕ simple with $0 \le \phi \le f$. Thus $\int f = \sup_\phi \int \phi \le \lim \int f_k$, which is what we wanted to show. $\square$

Proposition: $f \in L^+$ and $\int f = 0 \implies f = 0$ a.e.

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

Chebyshev's Inequality: If $f \in L^+$, then $m(\{x \ni f(x) \geq \alpha\}) \geq \frac{1}{\alpha} \int f \forall \alpha$.



Proof:

Just note that $\alpha m(\{x \ni f(x) > \alpha\}) = \int \alpha \chi_{f \geq \alpha} \leq \int f$. $\square$

Proof of Proposition:

This is obvious for simple functions. If $f = \sum a_i \chi_{E_i}$ in standard representation, then $\int f = \sum a_j m(E_j) = 0 \iff$ either $a_j = 0$ for all j, or $a_j \neq 0$ and $m(E_j) = 0$. So it only disagrees with zero on a measure zero set.

In general, suppose $f = 0$ a.e., noting that $\phi \leq f = 0 \implies \phi = 0$ a.e. and thus $\int phi = 0$ by the previous case. But then $\int f = \sup_\phi \int \phi = 0$ a.e., and we're done.

Now suppose $\int f = 0$ a.e., now we'll apply Chebyshev. This says that $m(\{x \ni f(x) \geq \alpha\}) = 0$ for any $\alpha \geq 0$. But then $m(\{x \ni f(x) > 0\}) = m(\bigcup_n \left\{x \ni f(x) > \frac{1}{n}\right\}) = \lim m(\text{stuff}) = \lim 0 = 0$. $\square$

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

Example:

- $f_k = k\chi_{[0, \frac{1}{k}]}$. 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!)

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 for a $\leq$.

Proof of Fatou: We have $\liminf_k f_k = \lim_k g_k$ where $g_k = \inf_{n\geq k} f_n$. Note that $g_k \nearrow \liminf_k f_k$, we can apply MCT. So

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

noting that $g_k \leq f_k \implies \int g_k \leq \int f_k$. $\square$

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

13 Tuesday Lecture

Two main convergence theorems: define $L^+ = \{f : \mathbb{R}^n \to [0,\infty] \ni f \text{ is measurable.}\}$. 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 $\int f_n \leq M$ uniformly, then $\int f \leq M$ (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 $\limsup \int f_n \leq \int f$ as well, we're done. Since integrals satisfy monotonicity, $f_n \leq f \implies \int f_n \leq \int f$ (a.e. But just applying order-limit laws, we have $\limsup \int f_n \leq \int f$ as well. $\square$

13.1 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 $\int |f| < \infty$. 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 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 $f = \text{Im} f + i\text{Re} f$, and define

$$\int f = \int \text{Re} f + i \int \text{Im} f.$$

> 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. Note that 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 $h = f + g \implies h_+ + h_- = f_+ + f_- + g_+ + g_-$, then maneuver things so that everything is positive and *then* take integrals.

We can provisionally define

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

where we'd like to define $\|f\|_1 = \int |f|$. But this only gives 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: If $f \in L^1$, then $|\int f| \leq \int |f|$.

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

$$\left| \int f \right| = \left| \int f_+ - \int f_- \right| \leq \left| \int f_+ \right| + \left| \int f_- \right| = \int f_+ + \int f_- = \int f_+ + f_- = \int |f|$$

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

$$\frac{(\int f)^*}{|\int f|} \int f \coloneqq \alpha \int f = \int \alpha f = \int \mathrm{Re}(\alpha f) + i \int \mathrm{Im}(\alpha f)$$

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

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

(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 $\|f\|_1 = \int_X |f|$.

We have

$$\int |f| = 0 \ a.e. \iff f = 0 \ a.e.$$

> Exercise: prove the above for L^+ functions.

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

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): If $\{f_n\} \in L^1$, $f_n \to f$ a.e., and $|f_n| \le g$ 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 $f_n \le g$ for all n, then $f \le g$. So $\int |f| \le \int |g| < \infty$, and $f \in L^1$.

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

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

$$\begin{aligned}
0 \le |f_n - f| \le |f|_n + |f| \le g + |f| \\
\implies g + |f| - |f_n - f| \ge 0 \\
\text{(This is where we'll apply Fatou.)} \\
\implies \int \liminf (g + |f| - |f_n - f|) \le \liminf \int (g + |f| - |f_n - f|) \\
\implies \int (g + |f|) \le \liminf \int (g + |f|) - \liminf \int |f_n - f| \\
\implies \int (g + |f|) \le \int (g + |f|) + \limsup \int |f_n - f| \\
\implies 0 \le \limsup \int |f_n - f|. \quad \square
\end{aligned}$$

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

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

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

Showing differentiability can be done by looking at

$$\lim \frac{|F(t_0) - F(t)|}{t - t_0} \le \lim \int \frac{f(x,t) - f(x,t_0)}{t - t_0}\, dx \le \int \lim \frac{f(x,t) - f(x,t_0)}{t - t_0}\, dx = \int f'(x,t_0)\, dx.$$

14 Lecture Thursday

For any measurable $X \subseteq \mathbb{R}^n$, we defined $L^1(X) = \{f : X \to \mathbb{C} \text{ measurable } \ni \int_X |f| < \infty\} / \sim$ 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.

This is a normed vector space with $\|f\|_{L^1(X)} \coloneqq \int_X |f|$.

> 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 $\|f_n - f\|_1 \to 0$.

Recall

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

- MCT: If $f_n \in L^+$ (important!!) and $f_n \nearrow f$ a.e. then $\lim \int f_n = \int f$.
  - Corollary: $\sum \int f_n = \int \sum f_n$.
- If $f_n \in L^1$ and $f_n \to f$ a.e. *and* $|f|_n \leq g$ with $g \in L^1$, then $\lim \int f_n = \int f$.
  - Stronger statement $f_n \xrightarrow{L^1} f$, i.e. $\int |f_n - f| \to 0$. The above only gives $|\int f_n - f| \to 0$. This follows because $\lim \int |f_n - f| =_{DCT} \int \lim |f_n - f| \to 0$, since $|f_n - f| \leq 2g$.

Theorem: If $f_n \in L^1, \sum \int |f|_n < \infty$, then $\sum f_n$ converges to an L^1 function and $\sum \int f_n = \int \sum f_n$.

> Note that uniform convergences is stronger than pointwise is stronger than a.e. convergence, and we should think of convergence in norm as weaker than all of these (although it is not 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, $\sum |f|_n \in L^1 \implies \sum |f_n(x)|$ is finite 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?

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

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

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

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.

14.1 Different Notions of Convergence

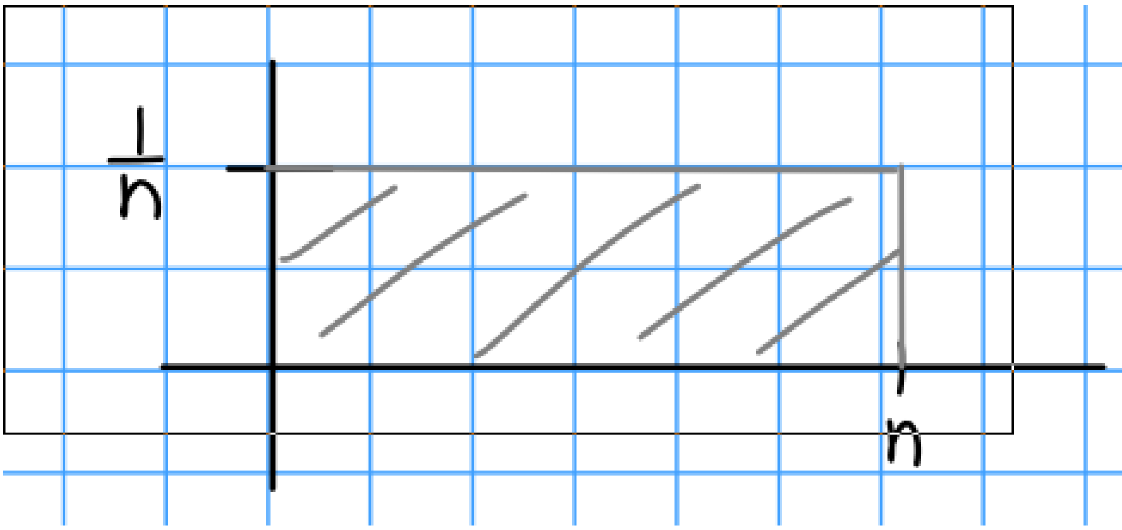
Note that $f_n \to f$ can mean many things:

1. Uniform: $f_n \rightrightarrows f : \forall \varepsilon\ \exists N \ni\ n \geq N \implies |f_N(x) - f(x)| < \varepsilon \quad \forall x$.
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: $\|f_n - f\|_1 = \int |f_n(x) - f(x)| \to 0$.

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.

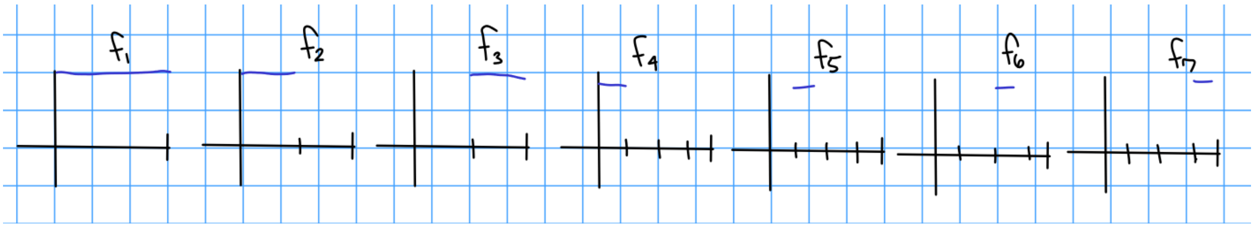


Figure 4: Image

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 $f_{n_k} = \sum^k g_j \to f$ a.e. We know it converges to *some* L^1 function, but limits are unique, so this is actually the original f. $\square$

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.

1. Find a candidate limit f
2. Show that the f_n actually converge to this candidate f
3. Show that f is in L^1.

Proof: Suppose f_n is Cauchy. From the previous proof, we know a subsequence in L^1 converges to some limit f in L^1. So call this f 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

$$\|f_n - f\|_1 \leq \|f_n - f_{n_k}\|_1 + \|f_{n_k} - f\|_1 \leq 2^{-k} + \varepsilon/2 \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

$\square$

15 Tuesday October 1

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

"$\mathbb{R}$ is complete" $\iff$ "$\sum |x|_n$"∞ $\implies$ $\sum x_n < \infty$".

Proof: $\implies$: A reminder of why the second statement holds: Suppose $\mathbb{R}$ is complete and $\sum |x|_n < \infty$. Let $S_N = \sum_{i=1}^N x_n$. Then if $N > M$, then $|S_N - S_M| \leq \sum_{i=M+1}^N |x|_n \to 0$.

$\impliedby$: 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, whose convergence implies convergence of our original series.

Choose a subsequence $n_1 \leq n_2 \leq \cdots$ such that $|x_n - x_m| < 2^{-j}$ if $n, m \geq n_j$. Let $y_1 = x_{n_1}$ and $y_j = x_{n_j} - x_{n_{j-1}}$ for $j > 1$. Then

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

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

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

as long as $n > n_k$ and k is sufficiently large. $\square$

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

"X is complete" $\iff$ "$\sum \|x\|_n$"∞ $\implies$ $\sum x_n < \infty$".

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

> Questions like the following show up on exams.

15.1 Translation and Dilation Invariance of the Lebesgue Integral

Define a translation $\tau_h(x) = x+h$, then $\tau f(x) = f(x-h)$ for all $h \in \mathbb{R}^\times$. Similarly, define a dilation $f_\delta(x) = \delta^{-n} f(\delta^{-1}x)$ for all $\delta > 0$.

Theorem: Given $f \in L^1$, then

1. $\tau_h f \in L^1$ and $\int \tau_h f = \int f$ (i.e. $\int_E f(x-h) = \int_{E+h} f$), and
2. $f_\delta \in L^1$ and $\int f_\delta = \int f$ (i.e. $\delta^{-n} \int f(\delta^{-1}x) = \int f(y)$)

Proof: First verify 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.

> Once you know something for simple functions, you can often apply MCT to get it for L^+ functions as well.

If $f \in L^+$ then there exists a sequence of simple functions $\{\phi_k\} \nearrow f$. 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 $\{\int \tau_h \phi_k\} \{\int \phi_k\}$, 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 $f = \text{Re}(f) + i\text{Im}(f)$.

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

$$G_p = \sum_{j=1}^n \sup\left\{f(x) \ni x \in [t_k, t_{j+1}]\right\} \chi_{[t_j,t_{j+1}]} g_p = \sum_{j=1}^n \inf\left\{f(x) \ni x \in [t_k, t_{j+1}]\right\} \chi_{[t_j,t_{j+1}]}.$$

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 $G_P \leq f \leq g_P$, and so $G \leq f \leq g$ as well.

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 and $\int G_{P_k} \to \int G$ and $\int g_{P_k} \to \int g$ and thus $\int g\ intG = \int_{\mathcal{R}} f$.

Since $\int G = \int g \implies \int (G-g) = 0 \implies G = g$ a.e., we have $f = G$ a.e. But G is a sequence of measurable function, and so f is measurable. Moreover, $\int G = \int f$. But $\int G = \int_{\mathcal{R}} f$ as well, so the two integrals agree. $\square$

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

> Examples next time: continuous functions with compact support are dense in L^1, which is a version of 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 $\int f - \int g < \varepsilon$.

16 Thursday October 3

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

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

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

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

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

$\impliedby$: Suppose $\int_E f = 0 \forall E \in \mathcal{M}$ 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 $E = \{x \ni f^+ > 0\}$, then $m(E) > 0$. Then $\int_E f^+ > 0$, since $\chi_E f^+ \neq 0$ a.e. and we know $f^+ \in L^+$ so $f^+ = 0$ a.e. iff $\int f = 0$. But then $\int_E f > 0$, since $\text{support}(f^+) \bigcap \text{support}(f^-) = \{x \ni f(x) = 0\}$, so $f^- = 0$ on E. $\square$

16.1 Approximation Theorems

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

$$\forall \varepsilon > 0 \;\&\; f \in L^1 \; \exists g \in \mathcal{C} \ni \|f - g\|_1 < \varepsilon.$$

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. (Remark) 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 $\int |\phi_k - f| < \varepsilon$ 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 there exists a simple function $\phi = \sum_{j=1}^N a_j \chi_{E_j}$ in standard representation, where $a_j \neq 0$ and the E_j are disjoint, with $\int |f - \phi| < \varepsilon$.

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

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

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

This approximation theorem yields some nice proofs:

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

1. "Small tails": $\exists N$ such that $\int_{\|x\| \geq N} |f| < \varepsilon$ (take $f_N = f\chi_{B(N)} \nearrow f$), and
2. "Absolute continuity": There exists a $\delta > 0$ such that $m(E) < \delta$ implies $\int_E |F| < \varepsilon$ (take $f_N = f\chi_{f(x) \leq N}$).

Prove something for L^1: 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 $E = \{x \ni |x| \geq N\}$. 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 a.e. Then $\int_E |g| \leq Mm(E) < \varepsilon$. So set $\delta = \varepsilon/M$, then if $m(E) < \delta$ then

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

Common qual problem: prove the following theorem. Note that DCT doesn't quite work!

Theorem (Continuity in L^1) Let $f \in L^1$, then $\lim_{h \to 0} \int |f(x+h) - f(x)| = 0$.

Proof: Let $\varepsilon > 0$. Then choose g such that $\int |f(x))g(x)| < \varepsilon$. By translation invariance, $\int |f(x+h) - g(x+h)| < \varepsilon$.

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 K such that $|g(x+h) - g(x)| = 0$ for all $x \in K^c$ and $|h| \leq 1$. But then $\int_K |g(x+h) - g(x)| \leq \varepsilon \int_K 1 \approx \varepsilon$.

Note that $\operatorname{supp}(F) = \overline{\{x \ni f(x) \neq 0\}}$.

17 Tuesday, October 8

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

17.1 Fubini and Tonelli

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

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. $f_x(y) \coloneqq f(x,y)$ is measurable as a function of y in $\mathbb{R}^{n_2}$

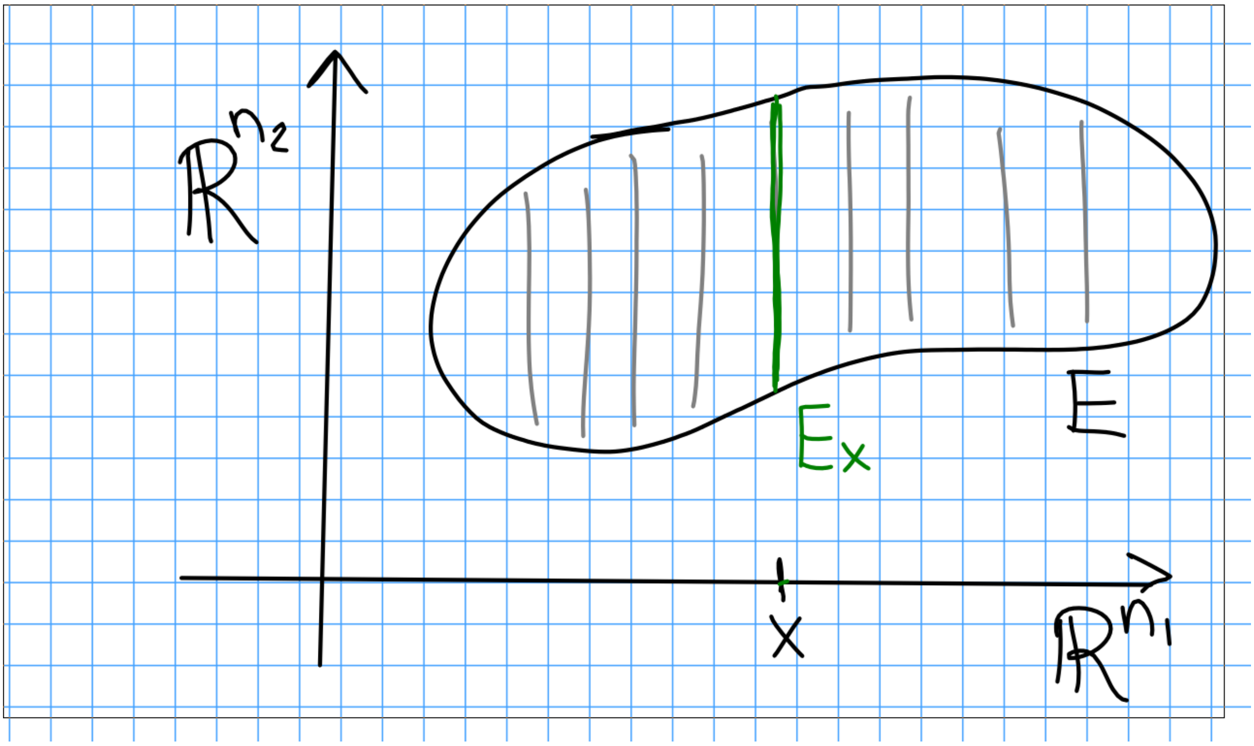


Figure 5: Image

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 equals $\int_{\mathbb{R}^n} f$.

Corollary: If $E \subset \mathcal{M}\ (\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$, then for a.e. $x \in \mathbb{R}^{n_1}$, the slice $E_x := \{y \in \mathbb{R}^{n_2} \ni (x,y) \in E\}$. 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; it is possible for every slice to be measurable while E itself is not. Take $E = \mathcal{N} \times I$, 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. Necessary in the context of convolution, since $f(x-y)$ will need to be measurable in both variables in order to apply Tonelli.

17.2 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

Application: Area under graph

Suppose $f \geq 0$ on $\mathbb{R}^n$, with no assumption of measurability. Consider defining the "area under the graph" as $\mathcal{A} := \{x, y) \in \mathbb{R}^n \times \mathbb{R} \ni 0 \leq y \leq f(x)\}$. Then

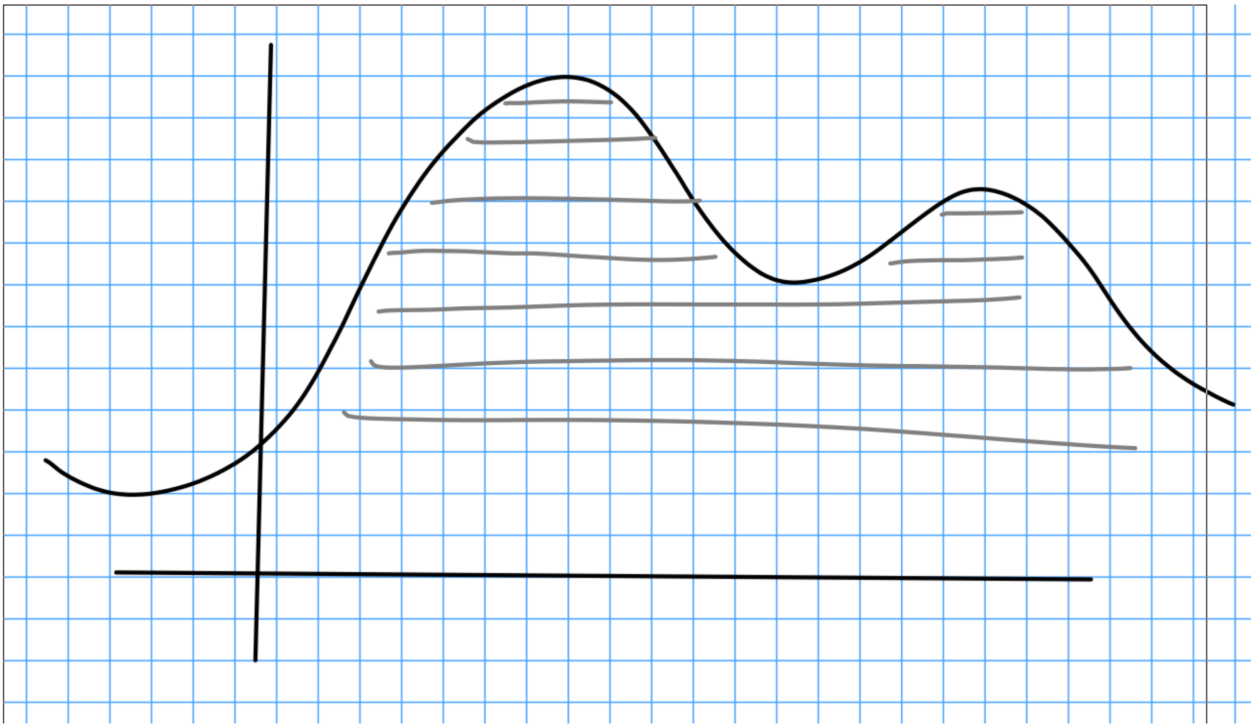


Figure 6: Image

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 $m(\mathcal{A}) = \int_{\mathbb{R}^n} f = \int_0^\infty m(\{x \in \mathbb{R}^n \ni f(x) \geq y\})\, dy$

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 $\mathcal{A}_x = \{y \in \mathbb{R} \ni (x,y) \in \mathcal{A}\} = [0, f(x)]$. Then $m(\mathcal{A}_x) = f(x)$, so $x \mapsto \mathcal{A}_x$ is a measurable function of x and $m(\mathcal{A}) = \int f(x)\, dx$.

Repeating this argument with y slices instead, for almost every $y \in \mathbb{R}$ we have $\mathcal{A}_y = \{x \in \mathbb{R}^n \ni (x,y) \in \mathcal{A}\} = \{x \in \mathbb{R}^n \ni f(x) \geq y \geq 0\}$, which is a measurable subset of $\mathbb{R}^n$. So it makes sense to integrate it, and thus $m(\mathcal{A}) = \int m(\mathcal{A}_y)\, dy = \int_0^y m(\{x \in \mathbb{R}^n \ni f(x) \geq y\})\, dy$. $\square$

Alternatively,

$$\begin{aligned}
\int_0^\infty m(\{x \in \mathbb{R}^n \ni f(x) \geq y\}) &= \int_{\mathbb{R}} \int_{\mathbb{R}^n} \chi_{S := \{x \in \mathbb{R}^n \ni f(x) \geq y \geq 0\}} \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}} \chi_S = \int_{\mathbb{R}^n} \int_0^{f(x)} dy\, dx \\
&= \int_{\mathbb{R}^n} f(x)\, dx.
\end{aligned}$$

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, the set $S_a = \{(x,y) \ni F(x,y) \geq a\} \in \mathcal{M}(\mathbb{R}^{n+1})$.

But $S_a = \{x \in \mathbb{R}^n \ni f(x) > a\} \times \mathbb{R}$, but this is the cylinder on a measurable set, which we will show 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 $M(\overline{Q}_i) = 2km(Q_i)$, which we can pull out of $\sum \left|\overline{Q}_i\right|$.

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. $\int_{\mathbb{R}^{n_2}} f(x, y)\ dy$ 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.

> Tonelli: Non-negative and measurable allows switching integrals, Fubini says if they're measurable then you can switch and the integrals are equal (and the integrals are finite). Fubini/Tonelli will allow switching *not* non-negative integrands.

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 $f \in L^1(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$ by Tonelli on $|f(x, y)|$.

Then by Fubini, $\int f$ is equal to either iterated integral.

> If any iterated integral is finite, then they all are.

Recall that $\sum \in f_n = \int \sum f_n$ is true exactly when

1. $f_n \geq 0$, and
2. $\sum \int |f_n| < \infty$.

18 Tuesday October 15

> Strong version of DCT: allows you to deduce L^1 converges from a.e. convergence. Note that otherwise, this are incomparable!

Proof of Fubini: see book. Essentially uses MCT a number of times and reduces to the case of cubes that possibly include boundaries.

19 Thursday October 17

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: Essentially use Fubini, and truncate domain/range with a k-ball. See proof of case-jumping lemma.

> We'll never need to dig into the proof of this, but there will always be a question related to applying it.

> Also, see notes for further detail.

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)
- $f \leq 0$ 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 $\{E \ni T^{-1}(E) \in \mathcal{B}\}$ is in fact a σ-algebra that contains all open sets.)

> Exercise: Run through this argument.

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 $m(N) = 0$. Then $T^{-1}(f^{-1}(G)) = T^{-1}(H) \bigcup T^{-1}(N)$. But by the first part, $T^{-1}(H)$ is still Borel, and $T^{-1}(N)$ is still null. But then $f \circ T$ is measurable.

> Note that this kind of thing usually works – just establish something Borel sets, then use this characterization to extend it to Lebesgue.

20 Tuesday October 22

20.1 Convolution

Recall:

- $C_c \hookrightarrow L^1$ is dense.
- Continuity in L^1: if $f \in L^1$,then $\|\tau f - f\| \to 0$, i.e. $\lim_{y \to 0} \int |f(x+y) - f(x)|\ dx = 0$.
- If $f \in L^1$, then for any $\varepsilon > 0$,
  - There exists a δ such that for all E such that $m(E) \leq \delta$, we have $\int_E |f| < \varepsilon$ (small tails).
  - There exists an N such that $\int_{\{\|x\| \geq N\}} |f| < \varepsilon$. (Note that $|f(x)| < \varepsilon \forall \|x\| \geq N$ 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.

Some remarks:

- There are conditions on f, g which can guarantee that $f \star g$ exists.
- If for some x, the function $y \mapsto f(x-y)g(y)$ is measurable, then $y \mapsto f(y)g(x-y)$ 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 $f \star g = g \star f$.

Theorem 1:

a. If $f \in L^1$ and g is bounded, then $f \star g$ is bounded *and* uniformly continuous.
b. If $f, g \in L^1$ and f, g are both bounded, then $\lim_{|x| \to \infty} f \star g(x) = 0$.

Note that (b) follows if $f \star g$ were uniformly continuous *and* integrable, but we don't necessarily need integrability for this result.

> Note that it is possible to pointwise multiply 2 integrable functions and get something not integrable – consider f^2 where $f(x) = \frac{1}{\sqrt{x}}\chi_{[0,1]}$.

Theorem 2: If $f, g \in L^1$, then $\|f \star g\|_1 \leq \|f\|_1 \|g\|_1$, and equality is attained if $f, g \geq 0$. (i.e. $\int |f \star g| \leq \int |f| \int |g|$.)

Corollary: If g is additionally bounded, then $\lim_{|x| \to \infty} f \star g(x) = 0$.

Theorem 3: If $f \in L^1$ and g *is differentiable* with $g, \frac{\partial g}{\partial x_1}, \cdots, \frac{\partial g}{\partial x_n}$ all bounded, then $f \star g \in C^1$ and $\frac{\partial}{\partial x_j}(f \star g) = f \star (\frac{\partial}{\partial x_j} g)$.

Corollary: If $f \in L^1$ and $g \in C_c^\infty \implies f \star g \in C^\infty$ and $\lim_{|x| \to \infty} f \star g(x) = 0$. 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 obtain compact support after this convolution. See the following picture, which looks similar for any fixed x – particularly any large x.

Proof of Theorem 1:

For (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

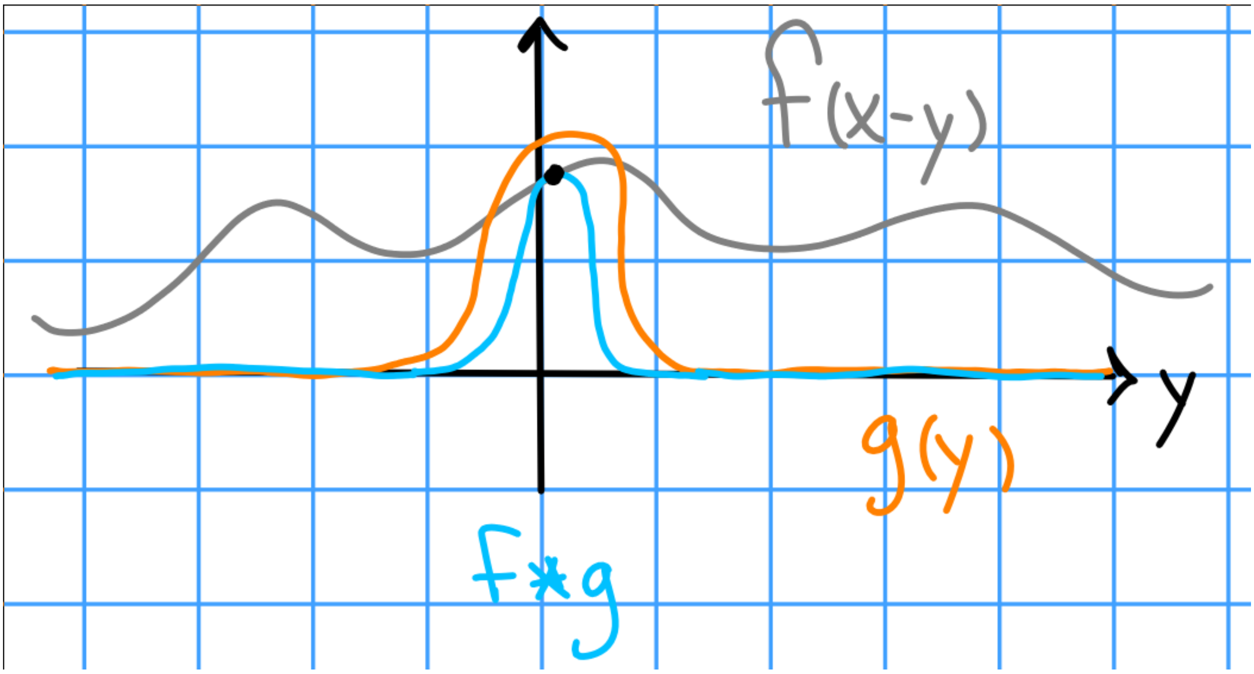


Figure 7: Image

$$
\begin{aligned}
|f \star g(x+h) - f \star g| &= \left| \int f(x+h-y)g(y)\ dy - \int f(x-y)g(y)\ dy \right| \\
&\leq \int |f(x+h-y) - f(x-y)||g(y)|\ dy \\
&\leq M \int |f(z+h) - f(z)|\ dz \to 0.
\end{aligned}
$$

For (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 $f, g \in L^1$, we have $h(x,y) := f(x-y)g(y)$ is measurable on $\mathbb{R}^n \times \mathbb{R}^n$. To see that f is in fact measurable, just define $F(x-y, y) = f(x-y)$ by taking the cylinder, then just let $T = [1, -1; 0, 1]$ so $T(x,y) = (x-y, y)$, so $\cdots = F \circ t(x,y)$.

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 a.e. 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 $f, g \geq 0$.

21 Thursday October 25?

Todo

22 Tuesday October 29

22.1 Approximations of the Identity

Theorem: Let $\phi \in L^1$ and $\int \phi = 1$. Then

- If f is bounded and uniformly continuous, then $f * \phi_t \to f$ uniformly where $\phi_t(x) := \frac{1}{t^n}\phi(\frac{x}{t})$
- If $f \in L^1$, then $f * \phi_t \to f$ in L_1.

Applications:

1. Theorem: $C_c^\infty \subset L^1$ is dense, i.e. $\forall \varepsilon > 0$ and for all $f \in L^1$, there exists a $g \in C_c^\infty$ such that $\|f - g\|_1 < g$.

Proof: Since $C - c$ is dense in L^1, it suffices to show the following:

$$\forall \varepsilon > 0 \& h \in C_c, \quad \exists g \in C_c^\infty \ni \|h - g\|_1 < \varepsilon.$$

Let $\phi \in C_c^\infty$ be arbitrary where $\int \phi = 1$ (which exist!). Then $\|h * \phi_t - h\|_1 < \varepsilon$ for t small enough. It remains to show that $h * \phi_t \in C_c^\infty$.

- This is smooth, because of theorem 3 regarding convolution applied infinitely many times.
- It is 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 $|x| > 2N$, then since $|x| \leq |x + y| + |y|$, so either $|x - y| \geq 2N$ or $|y| \geq N$. But then $h * \phi_t(x) = \int h(x - y)\phi_t(y)\, dy$, and by the previous statement, at least one term in the integrand is zero, so the integral is zero and $h * \phi_t$ is also compact supported.

2. Theorem (Weierstrass Approximation):

A function can be *uniformly* approximated by a polynomial on any closed interval, i.e.

$$\forall \varepsilon > 0,\ f \in C([a, b]), \quad \exists \text{ a polynomial } P \ni |f(x) - P(x)| < \varepsilon \quad \forall x \in [a, b].$$

Proof: Let g be a continuous function on $[-M, M] \supseteq [a, b]$ such that $g|_{[a,b]} = f$. 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 $|g * \phi_t(x) - g(x)| < \varepsilon \forall x \in [-M, M]$. Note that $\phi_t(x) = \frac{1}{t} e^{-\pi x^2/t^2}$, and Maclaurin expand to obtain $\frac{1}{t}\sum_{n=0}^\infty \frac{(-1)^n \pi^n x^{2n}}{t^{2n} n!}$.

Note that the Maclaurin series will converge uniformly on compact sets!

By uniform convergence of this series, we can truncate the series to bound the difference by, say, $\varepsilon / \|g\|_1$. Let $Q(x)$ be the truncated series. Then

$$|g * \phi_t(x) - g * Q(x)| \leq |g * (\phi_t - Q)(x)| \leq \|g\| \|p_t(x) - Q(x)\|_\infty < \varepsilon,$$

(where $\|f\|_\infty = \sup_{x \in [a,b]} |f(x)|$) and $g * Q(x)$ is a polynomial. $\square$

22.2 Fourier Transform on $\mathbb{R}^n$

Given $f \in L^1$, we defined the Fourier transform of f by

$$\hat{f}(\xi) = \int f(x) \exp(-2\pi i x \cdot \xi)\, dx.$$

Some things we know about the Fourier transform:

- $f \in L^1 \implies \hat{f}$ is bounded and uniformly continuous (from an old homework)

- $\lim |\xi| \to \infty \hat{f}(\xi) = 0$, i.e. the Riemann-Lebesgue lemma (see current homework), i.e. $\hat{f}$ vanishes at infinity.

> Warning: $f \in L^1 \not\Longrightarrow \hat{f} \in L^1$ necessarily!

Theorem (Inversion Formula): If $f, \hat{f} \in L^1$ then

$$f(x) = \int \hat{f}(x) \exp(2\pi i x \cdot \xi)\ d\xi \quad \text{for a.e. } x,$$

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. Then the 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 $\left\{ f \ni f, \hat{f} \in L^1 \right\}$ is dense in L^1.
2. $f \in L^1$ and $\hat{f} = 0$ almost everywhere implies that $f = 0$ almost everywhere (using the Inversion formula)

Proof of Inversion Formula:

> Note: Fubini-Tonelli won't work here *directly.*

We'll have

$$f(x) = \int \int f(y) \exp(-2\pi i y \cdot \xi) \exp(2\pi i x \cdot \xi)\ dy\ d\xi,$$

which is patently not in $L^1(\mathbb{R}^{2n})$. So we'll introduce a "convergence factor" $\exp(-\pi t^2 |\xi|^2)$, which will make the integral swap result in something integrable, then take limits.

Important example (HW): If $g(x) = e^{-\pi|x|^2}$ then $\hat{g}(\xi) = e^{-\pi|\xi|^2}$. Note $g_t(x) = \frac{1}{t^n} \exp(-\pi|x|^2/t^2)$ is an approximation to the identity, and $\int g_t = 1$. By a HW exercise, have have $\hat{g}_t(\xi) = \hat{g}(t\xi) = \exp(-\pi t^2 |\xi|^2)$, which is exactly the convergence factor we're looking for. Moreover, $f * g_t \to f$ 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

$$\int \hat{f}(\xi)\exp(-\pi t^2|\xi|^2)\exp(2\pi i x\cdot\xi)\ d\xi i$$

$$\begin{aligned}
&:= \int \hat{f}(\xi)\phi(\xi) \qquad (= f * g_t(x) \to_{L_1} f) \\
&= \int f(y)\hat{\phi}(y)\ dy \\
&=_{DCT} \int \hat{f}(\xi)\exp(2\pi i x\cdot\xi)\ d\xi \quad \text{as } t \to 0.
\end{aligned}$$

where $\phi(\xi) = \exp(2\pi i x\cdot\xi)\hat{g}_t(\xi)$.

HW: Shows $\hat{\phi}(y) = \hat{\hat{g}}(y-x)$, and an easy consequence is that this is in fact $g_t(x-y)$.

But now one term is converging to $\int \hat{f}(\xi)\exp(2\pi i x\cdot\xi)\ d\xi$ as $t \to 0$ pointwise, and $f * g_t(x) \to f$ as $t \to 0$ in L_1. Then 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.

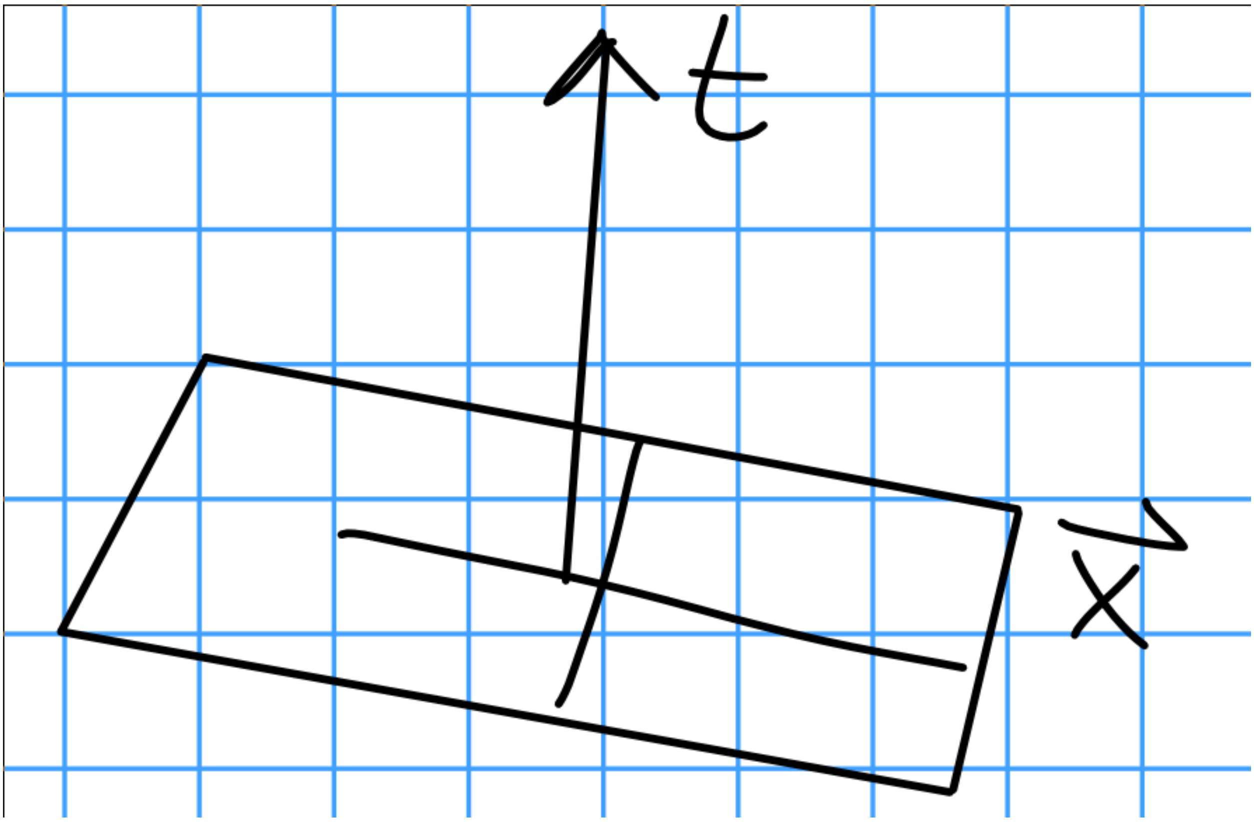
We thus obtain $f(x) = \int \hat{f}(\xi)\exp(2\pi i x\cdot\xi)\ d\xi$ almost everywhere.

23 Thursday October 31

Today: Some topics in PDEs.

23.1 The heat (diffusion) equation in the plane

Situation: 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

$$\frac{\partial u}{\partial t} = \frac{1}{4\pi}\left(\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}\right) := \frac{1}{4\pi}\Delta u$$
$$u(x, 0) = f(x).$$

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

$$\begin{aligned}\frac{\partial}{\partial t}\int_B u(x,t)\ dx &= \text{Flux across boundary}\\ &= \int_B \nabla \cdot \nabla_x u(x,t)\ dx \quad \text{by Green's/Divergence theorem}\\ &:= \int_B \Delta_x u(x,t)\ dx, .\end{aligned}$$

which is the heat equation.

23.1.1 Solution

We can use Fourier transforms to help solve these. Recall the identities:

- $\widehat{\frac{\partial}{\partial x_j} f}(\xi) = 2\pi i \xi_j \hat{f}(\xi)$.
- $\widehat{\frac{\partial^2}{\partial x_j^2} f}(\xi) = (2\pi i \xi_j)^2 \hat{f}(\xi) = -4\pi^2 \xi^2 \hat{f}(\xi)$.
- $\widehat{\Delta f}(\xi) = 4\pi^2 |\xi|^2 \hat{f}(\xi)$.

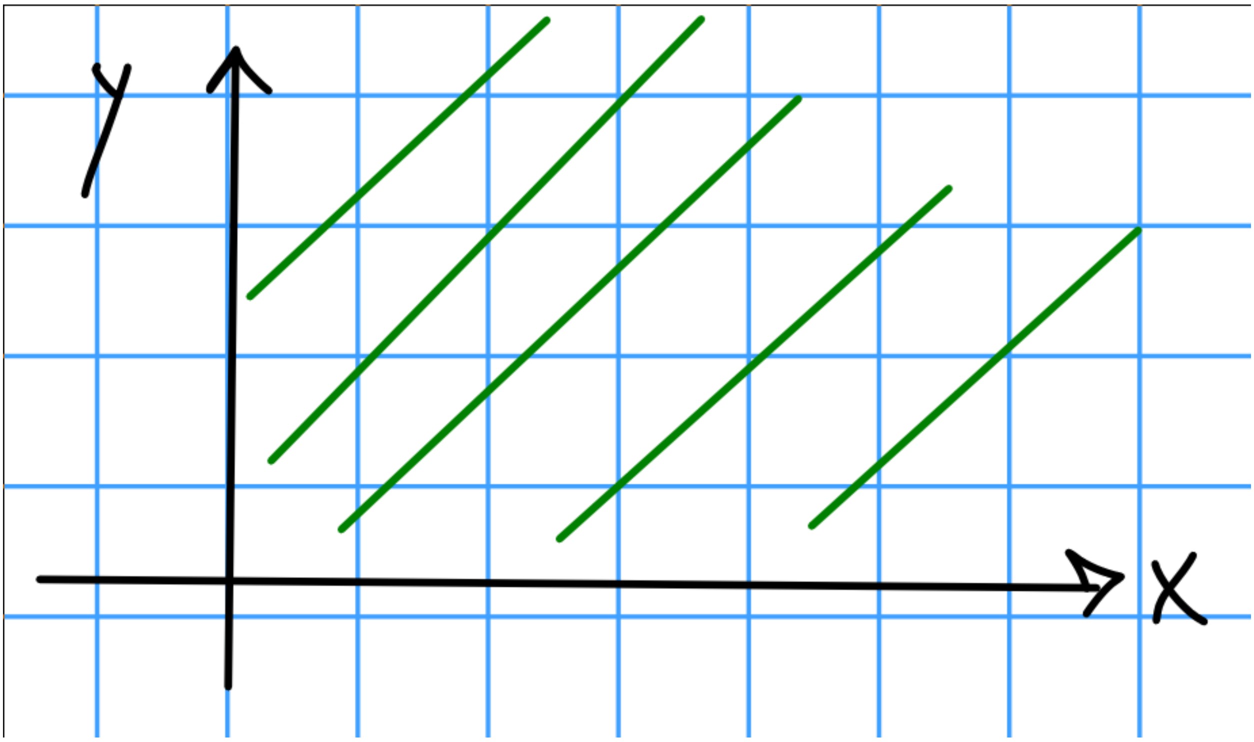
If we take the Fourier transform in the x variable, we get $\widehat{\frac{\partial u}{\partial t}} = \frac{\partial}{\partial t}\hat{u}(\xi, t) = -\pi |\xi|^2 \hat{u}(\xi, t)$. Then the boundary conditions become $\hat{u}(\xi, 0) = \xi f(\xi)$. But note that this is now a first order ODE!

This is easy to solve, we get $\hat{u}(\xi, t) = c(\xi) \exp(-\pi |\xi|^2 t) = \hat{f}(\xi) \exp(-\pi |\xi|^2 t)$.

But then $\exp(-\pi |\xi|^2 t) = \hat{G}(\sqrt{t}\xi)$ where $G(x) = \exp(-\pi |x|^2)$. 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 $u(x, t) = f * G_{\sqrt{t}}(x)$ where $G_{\sqrt{t}}(x) = \frac{1}{t^{n/2}} \exp(-\pi |x|^2 / t)$. 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.

23.2 Dirichlet problem in the upper half-plane

Situation:



We want to solve

$$\Delta u = 0$$
$$u(x, 0) = f(x).$$

23.2.1 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

$$\hat{u}(\xi, y) = A(\xi)\exp(-2\pi|\xi|y) + B(\xi)\exp(2\pi|\xi|y)$$

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 $\hat{u}(\xi, y) = \hat{f}(\xi)\hat{P}(y\xi) = \widehat{f * P_y}$ 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 involves a *second* derivative!

23.2.2 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 actually equivalent.

24 Tuesday: November ?

Today: Hilbert Spaces

> See notes on the webpage.

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 Schwarz inequality, i.e.

$$|\langle x,\ y\rangle| \leq \|x\|\|y\|,$$

with equality iff $x = \lambda y$.

We'll use the following computation often:

$$
\begin{aligned}
\|x+y\|^2 &= \langle x+y,\ x+y\rangle \\
&= \|x\|^2 + 2\mathrm{Re}\langle x,\ y\rangle + \|y\|^2 \\
&\leq \|x\|^2 + |\langle x,\ y\rangle| + \|y\|^2 \\
&\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \qquad \text{by Schwarz} \\
&= (\|x\| + \|y\|)^2.
\end{aligned}
$$

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:

- $\mathbb{C}^n$ with $\langle x,\ y\rangle = \sum x_j \overline{y_j}$
- $\ell^2(\mathbb{N})$ with $\langle x,\ y\rangle = \sum^\infty x_j y_j$ with $\sum |x_j| < \infty$ and similarly for y. Note that this is finite by AMGM, since $\sum x_i y_i \leq \frac{1}{2}(\sum x_i + \sum y_i) < \infty$ by assumption.
- $L^2(\mathbb{R}^n)$ with $\langle f,\ g\rangle = \int f\overline{g}$. This is also finite because $|f\overline{g}| \leq \frac{1}{2}(\int f + \int g)$.

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* $\lambda \in CC$, so $\langle x - \lambda y,\ x - \lambda y\rangle > 0$. This equals $\|x\|^2 - 2\overline{\lambda}\mathrm{Re}\langle x,\ y\rangle + |\lambda^2|\|y\|$. 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 $4|\langle x,\ y\rangle|^2 - 4\|y\|^2\|x\|^2$, which yields Cauchy-Schwarz.

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 $\|x \pm y\|^2 + \|x\|^2 \pm 2\mathrm{Re}\langle x,\ y\rangle + \|y\|^2$, so just add and the cross-terms will cancel.

Proposition (Pythagorean):

$$\langle x,\ y\rangle = 0 \implies \|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

In this situation, we say x, y are orthogonal.

Corollary: If $\{x_i\}$ are all pairwise orthogonal, then $\|\sum x_i\|^2 = \sum \|x_i\|^2$.

Orthonormal Sets:

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.

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:

$$\begin{aligned}
\|x - \sum a_n u_n\|^2 &= \|x\|^2 - 2\mathrm{Re}\sum \langle x,\ u_n\rangle a_n + \sum |a_n|^2 \\
&= \|x\|^2 - \sum |\langle x,\ u_n\rangle|^2 + \sum \left(|\langle x,\ u_n\rangle|^2 - 2\mathrm{Re}\langle x,\ u_n\rangle a_n + |a_n|^2\right) \\
&= \|x\|^2 - \sum |\langle x,\ u_n\rangle|^2 + |\langle x,\ u_n\rangle - a_n|^2 & \geq 0,
\end{aligned}$$

where equality is attained iff $a_n = \langle x,\ u_n\rangle = a_n$. So this is the best approximation.

> Hint: the pieces that are equalities are somehow easier to show – they necessarily involve direct computations.

But then

$$0 \leq \|x - \sum \langle x,\ u_n\rangle u_n\|^2 = \|x\|^2 - \sum |\langle x,\ u_n\rangle|^2,$$

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.

Theorem (Riesz-Fischer):

The map $x \xrightarrow{hat} \langle x,\ u_n\rangle \coloneqq \hat{x}(n)$ maps H onto ℓ^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 $n \in N$. Moreover, x can be chosen such that $\|x\| = \sqrt{\sum |a_n|^2}$.

> Note: this is not a bijection, there may not be a unique x, and a_n are referred to as the Fourier coefficient. Also note that if in fact $a_n = 0$ for all n implies $x = 0$, 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, which is also equivalent to something else.

25 Thursday November 7

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 (equation 1)

$$0 \leq \|x - \sum_{n=1}^{N} \langle x,\ u_n\rangle u_n\|^2 = \|x\|^2 - \sum_{n=1}^{N} |\langle x,\ u_n\rangle|^2 \forall N.$$

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
- $\lim_{N\to\infty} \|x - \sum_n^N \langle x,\ u_n\rangle\| = 0$, 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 $\exists x \in H$ such that $a_n = \langle x,\ u_n\rangle$ 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. $\langle y,\ u_n\rangle = 0 \forall n \implies y = 0$.

Proof; Let $S_N := \sum_{n=1}^N a_n u_n$. Then S_N is Cauchy, so

$$\begin{aligned}
\|S_N - S_M\|^2 &= \|\sum_{n=M+1}^{N} a_n u_n\|^2 \\
&= \sum_{n=M+1}^{N} \|a_n u_n\|^2 \qquad \text{by Pythagoras since the } u_n \text{ are orthogonal} \\
&= \sum_{n=M+1}^{n} |a_n| \to 0,
\end{aligned}$$

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 is large enough, in particular $N \geq n$, 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. $\square$

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 $\|x\|^2 = \langle x,\ x \rangle = \langle x,\ x - y \rangle \leq \|x\| x - y \leq \varepsilon \|x\| \to 0$. $\square$

> 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 $x = y$. $\square$

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:

- $\ell^2\mathbb{N}$: $u_n(k) = 1 \iff n = k$ and 0 otherwise, i.e. $\mathbf{e}_i$.
- $L^2([0,1])$: $e_n(x) \coloneqq e^{2\pi i n x}$. 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.

Sketch proof that $L^2([0,1])$ is complete:

Note that $L^2([0,1]) \subseteq L^1([0,1])$, since $f \in L^2 \implies \int_0^1 |f| 1\, dx \leq \sqrt{\int_0^1 |f|^2}$ by C-S. 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.

Now $\int \liminf_k \left| f_{n_j} - f_{n_k} \right|^2 \leq \liminf \int \left| f_{n_j} - f_{n_k} \right|$ by Fatou. But the LHS goes to $\int \left| f_{n_j} - f \right|$ and the RHS is $\|f_{n_j} - f_{n_k}\| \to 0$, so less than ε if j is big enough. So $f_{n_j} \to f$ in L^2 as $j \to \infty$, and thus $f_n \to f \in L^2$ as $n \to \infty$.

Unitary Maps:

Let $U : H_1 \to H_2$ such that $\langle Ux,\ Uy \rangle = \langle x,\ y \rangle$ (i.e. U preserves angles). Then $\|Ux\| = \|x\|$, i.e. U is an isometry. If U is surjective, this implication can be reversed.

For example, taking the Fourier transform yields $\sum \left| \hat{f}(u) \right|^2 = \|f\|_2^2 = \int |f|^2$, 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 $\ell^2\mathbb{N}$.

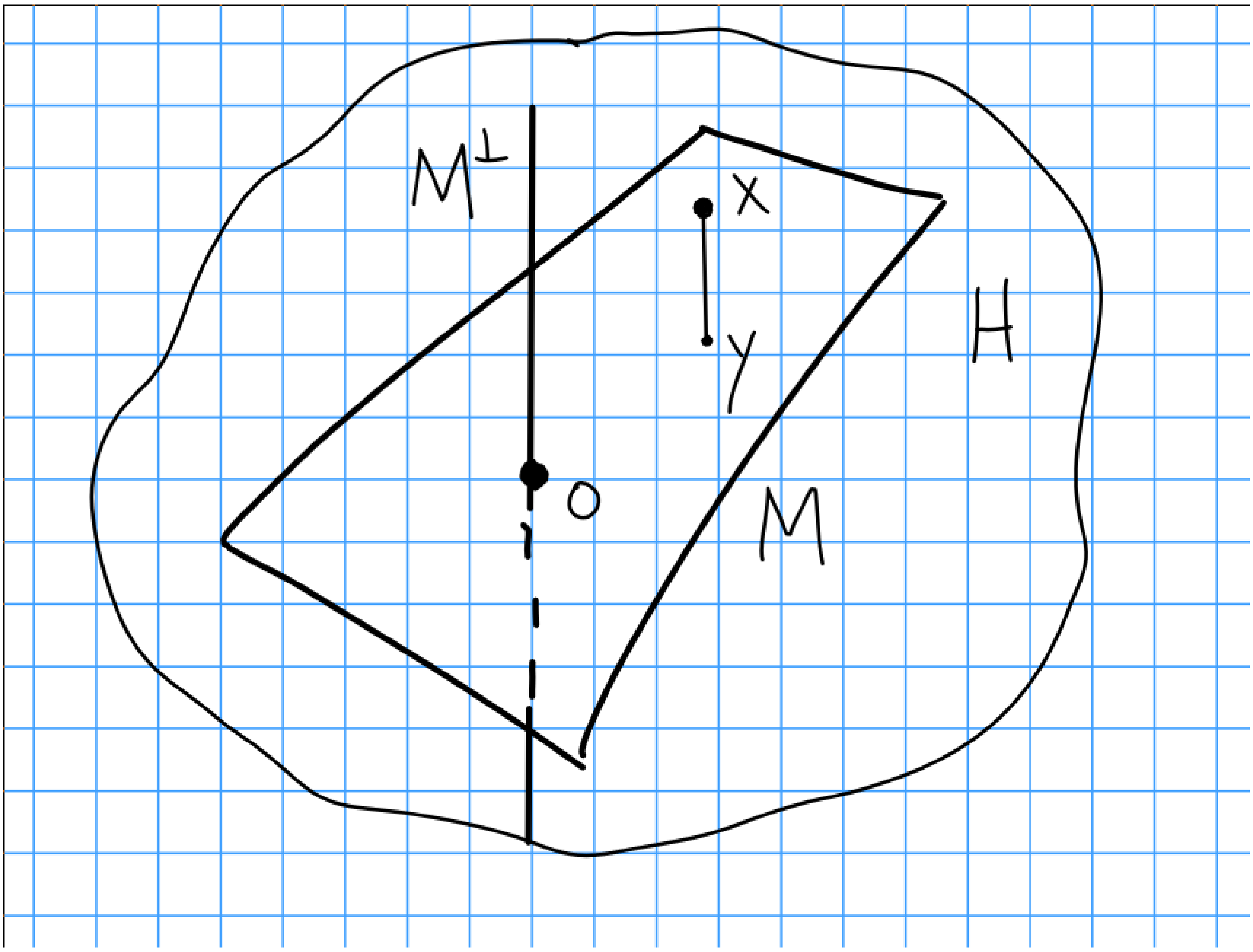


Figure 8: Image

Subspaces in Hilbert spaces don't have to be closed, but orthogonal complements are always closed! See homework problem.

26 Tuesday November 12

26.1 Closed Subspaces and Orthogonal Projections

Let H be a Hilbert space, then a subspace $M \subseteq H$ is *closed* if $x_n \to_H x$ 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 $M^\perp := \{x \in H \ni \langle x,\ y\rangle = 0\ \forall y \in M\}$ 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. $\|x - y\| = \inf_{y' \in M} \|x - y'\|$.
2. Defining $z := x - y$, 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. $x \in H \implies x = y + z, y \in M, z \in M^\perp$, and y, z are the unique elements in $M, M^\perp$ that are closest to x.

Proof of Lemma:

Part 1: Let $\delta := \inf_{y' \in M} \|x - y'\|$, 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 $\|x - y_n\| \to \delta$.

For the following parallelogram:

Then by the parallelogram law, we have

$$2(\|y_n - x\|^2 + \|y_m - x\|^2) = \|y_n - y_m\|^2 + \|y_n + y_m - 2x\|^2.$$

which yields

$$\begin{aligned}
\|y_n - y_m\|^2 &= 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\|\frac{1}{2}(y_n + y_m) - x\|^2 \\
&\leq 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\delta^2 \to 0,
\end{aligned}$$

since $\|y_n - x\|_H \to 0$ since $y_n \to_H x$.

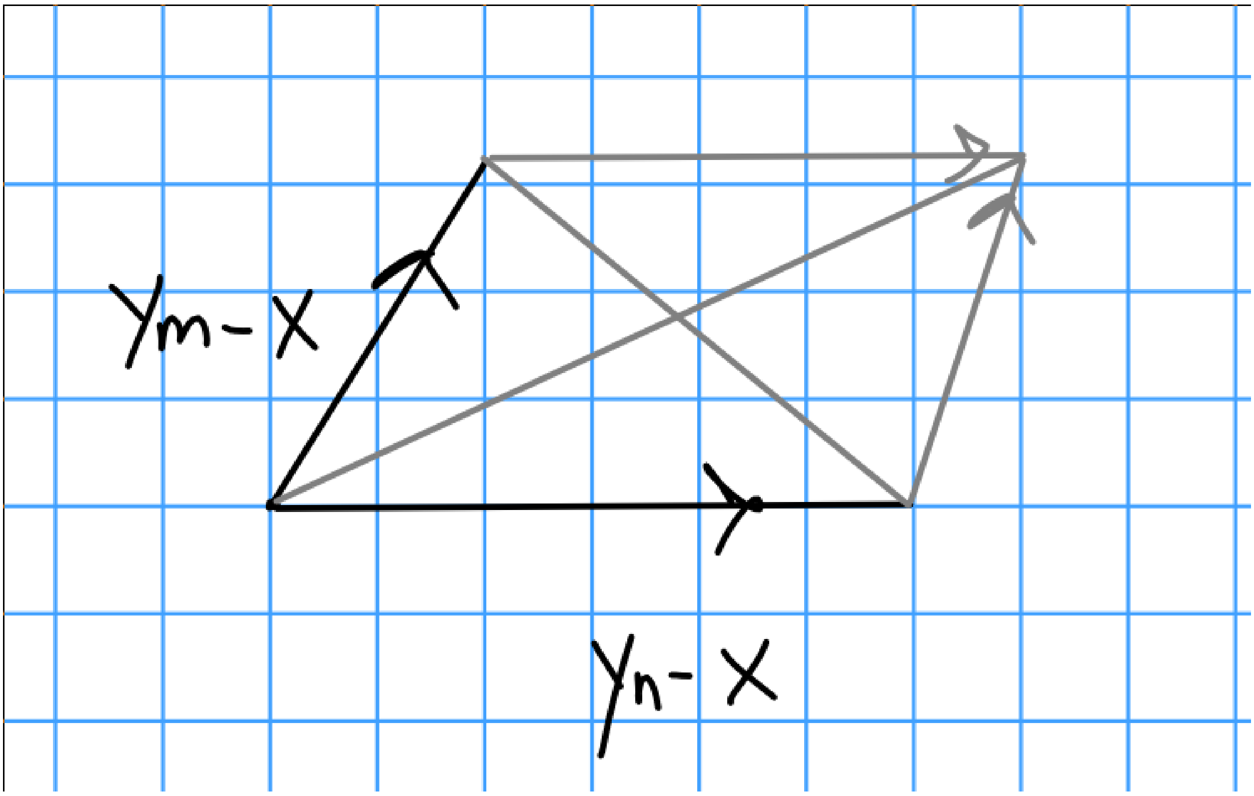


Figure 9: Image

It follows that $\{y_n\}$ is Cauchy in H, so $y_n \to_H y \in H$. 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.

Part 2: Let $u \in M$, we want to show that $\langle z, u \rangle = \langle x - y, u \rangle = 0$.

Wlog 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 things here, less so this proof.

26.2 Trigonometric Series

Theorem: Let $e_n(x) := e^{2\pi n x}$ 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 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:

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 $|f(x) - P(x)| < \varepsilon$ for all $x \in \Pi$.

Note that this measures closes in the *uniform* norm. We can relate these by $\|f(x) - P(x)\|_{L^2} \leq \|f(x) - P(x)\|_{\sup}$, 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

- $Q_k(x) \geq 0$ for all x, k,
- $\int_{-1/2}^{1/2} Q_k(x)\ dx = 1$ for all k,
- $\forall \delta > 0$, we have $Q_k(x) \to 0$ uniformly on $\Pi \setminus [-\delta, \delta]$.

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 on 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 $|f(x-y) - f(x)| < \varepsilon/2$ for all x in the δ ball. But this lets us break the integral into two pieces, $I = \int_{y\in B_\delta} \cdots\ dy + \int y \in B_\delta^c \cdots\ dy$, where the second term can be made smaller than $\varepsilon/2$ by taking k large enough.

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 cosine is decreasing on $[\delta, \frac{1}{2}]$, so $Q_k(x) \leq Q_k(\delta) = c_k \left(\frac{1+\cos(2\pi\delta)}{2}\right)^k$. 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.