Real Analysis and Measure Theory

James

January 6, 2020

Contents

<u>Midterm 1: February 27</u>
Topics:

(a) Lebesgue measure on n-dimensional Euclidean space (Chapter 2)

(b) Measurable functions (Chapter 3)

(c) Lebesgue integration, including the limit theorems (Fatou, MCT, DCT, UCT, BCT) (Chapter 4)

(d) Fubini's theorem (Chapter 4)

<u>Results of Midterm 1:</u>
Median: 47
Average: 51
Total Possible: 75

<u>Midterm 2: April 10</u>
Topics:

(a) Lebesgue Differentiation Theorem (Chapter 5)

(b) Differentiation of Monotone Functions (Chapter 5)

(c) Absolutely Continuous Functions (Chapter 5)

(d) L^p classes (including Banach and metric space properties) (Chapter 6)

(e) Hölder's inequality (Chapter 6)

(f) Jensen's inequality (Chapter 6)

(g) Convolutions (Chapter 4 Lecture 17, Chapter 6)

(h) Approximations of the identity (Chapter 6)

Chapter 1

Prerequisites

This is more of a list of things one should know reading this rather than an educational chapter.

Definition. For a point $x \in \mathbb{R}^n$ and $\epsilon > 0$ we define the **open ball of radius ϵ** to be

$$B(x, \epsilon) = B_\epsilon(x) = \{y \ : \ |x - y| < \epsilon\}.$$

The notation will be used interchangeably (the notation on the right being mine and the left being the books).

Definition. A point x of a set E is called an **interior point** of E if there exists a $\delta > 0$ such that $B(x, \delta) \subseteq E$. The collection of all interior points of E is called the **interior**, and is denoted by E°.

Definition. A point $x_0 \in \mathbb{R}^n$ is said to be a **limit point of a set E** if it is the limit point of a sequence of distinct points in E.

Definition. A point $x_0 \in E$ is said to be a **isolated point of a set E** if it is not the limit of any sequence in E outside of the trivial sequence. In other words, it is an isolated point if and only if there exists a $\delta > 0$ such that $|x - y| > \delta$ for all $y \in E \backslash \{x\}$.

Definition. A set is said to be **open** if $E = E^\circ$.

Definition. A set is said to be **closed** if E^C is open.

Remark 1. Open sets are closed under arbitrary union and countable intersection. DeMorgan's Laws give us the opposite for closed sets; that is, closed sets are closed under arbitrary intersection and countable union.

Definition. A union of a set E along with all of its limit points is called the **closure** of E, and is denoted by $\bar{E}$.

Definition. A function $f : X \to Y$ is said to be **continuous** if, for all open $V \subseteq Y$, we have that $f^{-1}(V)$ is open in X.

For Euclidean space, we define continuity if

$$\lim_{x \to x_0} f(x) = f(x_0).$$

Definition. If f is only defined on a set E containing x_0, $E \subseteq \mathbb{R}^n$, then f is said to be **continuous at x_0 relative to E** if $f(x_0)$ is finite and either x_0 is an isolated point or x_0 is a limit point of E and

$$\lim_{\substack{x \to x_0 \\ x \in E}} f(x) = f(x_0).$$

Definition. Let E be a set. Then we say that $\mathscr{F}$, a collection of sets, is a **cover** of E if

$$E \subseteq \bigcup_{K \in \mathscr{F}} K.$$

We say that $\mathscr{F}$ is an **open cover** if $K \in \mathscr{F}$ is open for all K.

Definition. Let E be a set. Then we say that E is **compact** if for every open cover $\{K_i\}_{i=1}^{\infty}$ of E we have that there exists a finite subcollection $\{K_i\}_{i=1}^{N}$ that covers E.

Theorem 1.1. (Heine-Borel Theorem) We have that a set $E \subseteq \mathbb{R}^n$ is compact if and only if it is closed and bounded.

Theorem 1.2. A set $E \subseteq \mathbb{R}^n$ is compact if and only if every sequence of points of E has a subsequence that converges to a point of E.

Lemma 1.1. If ϕ is a continuous function on a compact set E, we have that f is bounded.

Remark 2. (Open Sets in $\mathbb{R}^n$) For $n = 1$, every open set is a countable union of disjoint open intervals. For $n \geqslant 2$, every open set is a union of a countable collection of non-overlapping closed boxes.

Definition. A sequence $\{f_n\}_{n \in \mathbb{N}}$ of continuous functions on an interval $I = [a, b]$ is **uniformly bounded** if there is a number M such that

$$|f_n(x)| \leqslant M$$

for all n and for all $x \in I$.

Definition. A sequence $\{f_n\}_{n \in \mathbb{N}}$ of continuous functions is said to be **equicontinuous** if for every $\epsilon > 0$ and x there exists a $\delta > 0$ such that

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

whenever $|x - y| < \delta$ for all functions f_n in the sequence. Note that δ may depend on ϵ and x but not y nor n.

Theorem 1.3. (Arzela-Ascoli Theorem) If a sequence $\{f_n\}$ of continuous functions is bounded and equicontinuous, then it has a uniformly convergent subsequence.

Definition. A sequence of functions $\{f_n\}$ **converges pointwise** to f if for all $x \in E$, where E is the domain, we have

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

Definition. A sequence of functions $\{f_n\}$ **uniformly converges** to f if, for all $\epsilon > 0$, there exist an N such that for all $n \geqslant N$, we have

$$|f_n - f| < \epsilon.$$

Definition. We define the **supremum-norm** on a set S to be

$$||f||_\infty = ||f|| = \sup\{|f(x)| \ : \ x \in S\}.$$

Theorem 1.4. (Stone-Weierstrauss Theorem) Suppose f is a continuous real-valued function defined on an interval $[a, b]$. For every $\epsilon > 0$, there exists a polynomial p such that for all $x \in [a, b]$ we have $|f(x)-p(x)| < \epsilon$, or equivalently the supremum-norm $||f - p|| < \epsilon$.

Remark 3. In other words, for every continuous function ϕ on a compact set E we can construct a sequence of polynomials P_n which converge uniformly to ϕ on E.

Definition. A **partition** of an interval $I = [a, b]$ is a set $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ such that $x_0 = a$, $x_n = b$, and $x_i \leqslant x_{i+1}$ for all $0 \leqslant i \leqslant n-1$. We say that $\mathcal{P}$ is size $n+1$, and denote this by $|\mathcal{P}| = n+1$.

Definition. Given a bounded function $f : [a, b] \to \mathbb{R}$ and partition $\mathcal{P}$ of [a,b] of size $n+1$ with associated partitioning intervals I_i, we define the **upper Riemann sums of f with respect to $\mathcal{P}$** to be

$$U(f, [a, b], \mathcal{P}) = \sum_{i=1}^{n} M_i(f, \mathcal{P})\delta(I_i),$$

where

$$M_i(f, \mathcal{P}) = \sup_{x \in I_i} f(x),$$

$$I_i = [x_{i-1}, x_i],$$

$$\delta(I_i) = x_{i+1} - x_i.$$

Definition. In a similar setting to above, we define the **lower Riemann sums of f with respect to $\mathcal{P}$** to be

$$L(f, [a, b], \mathcal{P}) = \sum_{i=1}^{n} m_i(f, \mathcal{P})\delta(I_i),$$

where

$$m_i(f, \mathcal{P}) = \inf_{x \in I_i} f(x).$$

Definition. Given a bounded function $f : [a,b] \to \mathbb{R}$, we define the **upper Riemann integral of f on $[a,b]$** to be

$$\int_{[a,b]}^{+} f = \inf_{\mathcal{P}} U(f, [a,b], \mathcal{P}).$$

The **lower Riemann integral** is defined analogously;

$$\int_{[a,b]}^{-} f = \sup_{\mathcal{P}} L(f, [a,b], \mathcal{P}).$$

Definition. We say that a bounded function $f : [a,b] \to \mathbb{R}$ is **Riemann integrable** is

$$\int_{[a,b]}^{+} f = \int_{[a,b]}^{-} f.$$

If it is Riemann integrable, we define the **Riemann integral of f on $[a,b]$** to be

$$\int_{a,b} f = \int_{[a,b]}^{-} f = \int_{[a,b]}^{+} f.$$

Remark 4. If a function is continuous, it is Riemann integrable.

Remark 5. If a function is bounded and monotone on $[a,b]$, then f is Riemann integrable on $[a,b]$.

Definition. We define the **support** of a function to be the closure of the set of points where the function is non-zero.

Definition. We say that a function f has **compact support** if it is 0 outside of a compact set.

Definition. **Variation** is defined to be

$$V(f;[a,b]) = \sup_{\mathcal{P}} \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|,$$

where $\mathcal{P}$ is a partition of the interval $[a,b]$.

Definition. A function has **bounded variation** if $V(f;[a,b]) \leqslant M$ for some fixed M.

Definition. A **vector space** V over a field F is a set which is closed under vector addition and scalar multiplication.

Definition. A **norm space** is a vector space V equipped with a function $||\cdot|| : V \to [0,\infty)$ such that

(i) $||v|| = 0$ if and only if $v = 0$,

(ii) $||cv|| = |c|||v||$,

(iii) $||v + w|| \leqslant ||v|| + ||w||$.

Definition. A sequence $\{a_n\}$ is **Cauchy** if for all $\epsilon > 0$, there exists an N such that for all $n, m \geqslant N$ we have

$$||a_n - a_m|| < \epsilon.$$

Definition. A sequence $\{a_n\}$ converges if there exists an a such that for all $\epsilon > 0$ there exists an N such that for all $n \geqslant N$ we have

$$|a_n - a| < \epsilon.$$

Definition. A space V is **complete** if every Cauchy sequence converges. That is, if $\{a_n\}$ is Cauchy, then there is some $v \in V$ such that $a_n \to v$.

More will be added as needed.

Chapter 2

Measure on Euclidean Space

2.1 Lecture 1 (Outer Measure)

The first goal is to define some way of measuring volume of sets in $\mathbb{R}^n$. We first start with some notation. Throughout, $I = \{x \in \mathbb{R}^n \ : \ a_i \leqslant x_i \leqslant b_i \text{ for } i = 1, \ldots, n\}$ for some $a_i < b_i$; i.e. I will denote closed intervals. If $E \subseteq \mathbb{R}^n$ and $E \subseteq \bigcup_k I_k$ (countable), then we call $\{I_k\}$ a cover of E.

Definition. We define the **volume of a closed interval** to be

$$v(I) = \prod_i (b_i - a_i).$$

Definition. We define the **outer measure** or **exterior measure** of a set E to be

$$|E|_e = \inf_{\text{covers } \{I_k\} \text{ of } E} \sum_k v(I_k).$$

Intuitively, the exterior measure is just covering the set with smallest number of cubes possible and then adding up the volume of those cubes. This is *almost* the definition of measure we're going to use. The problem is that there are some sets where this measure is not any good; that is, we want the outer measure to be the same as the inner measure, and in some cases this will not happen. When it does, though, we will call this set **measurable**.

Theorem 2.1. (Properties of The Outer Measure)

(i) $|I|_e = v(I)$ for all closed intervals I.

(ii) If $E_1 \subseteq E_2$, then $|E_1|_e \leqslant |E_2|_e$.

(iii) If $\{E_k\}$ is a countable collection of sets, then

$$\left|\bigcup_k E_k\right|_e \leqslant \sum_k |E_k|_e.$$

Proof. (i) First, note that I is a cover of itself. Thus, we get $|I|_e \leqslant v(I)$. For the reverse, let $I \subseteq \bigcup_k I_k$. Let I_k^* be a "small blowup"; that is,

$$I_k \subseteq (I_k^*)^\circ \subseteq I_k^*$$

be such that

$$v(I_k^*) \leqslant (1+\epsilon)v(I_k), \quad \epsilon > 0.$$

Since I is compact,

$$I \subseteq \bigcup_{k=1}^N (I_k^*)^\circ \subseteq \bigcup_{k=1}^N I_k^*.$$

Now, $I \subseteq \bigcup_{k=1}^N I_k^*$ implies $v(I) \leqslant \sum_{k=1}^N v(I_k^*)$. Thus, we get

$$v(I) \leqslant \sum_{k=1}^N v(I_k^*) \leqslant (1+\epsilon)\sum_{k=1}^N v(I_k) \leqslant (1+\epsilon)\sum_k v(I_k).$$

Thus, $v(I) \leqslant \sum_k v(I_k)$ for any $I \subseteq \bigcup_k I_k$, and so we get $v(I) \leqslant |I|_e$.

(ii) This comes directly from the definition. Since $E_1 \subseteq E_2$, we get that all covers of E_2 also cover E_1, and so we get that the measure of E_2 can be at most the measure of E_1.

(iii) We cover E_k by $\bigcup_j I_j^{(k)}$ such that

$$\sum_j v\left(I_j^{(k)}\right) \leqslant |E_k|_e + \epsilon 2^{-k},$$

where $\epsilon > 0$ is arbitrary. Thus, we have that $\bigcup_k E_k$ is covered by $\bigcup_{j,k} I_j^{(k)}$. So

$$\left|\bigcup_k E_k\right| \leqslant \sum_{j,k} v\left(I_j^{(k)}\right) \leqslant \sum_k \left(|E_k|_e + \epsilon 2^{-k}\right) = \sum_k |E_k|_e + \epsilon.$$

Since ϵ was arbitrary, take the limit as $\epsilon \to 0$. This gives us the desired result.

Q.E.D

2.2 Lecture 2 (Lebesgue Measure)

We now want to briefly discuss the independence of outer measure from the choice of axis. When discussing outer measure originally, we covered our plane with boxes which have edges *parallel* to the axis. What if, however, we had diagonal axis? It should be that this gives us the same result. For notation purposes, we will denote things in alternative axis with a prime.

Theorem 2.2. For $E \subseteq \mathbb{R}^n$, we have

$$|E|_e = |E|'_e.$$

Proof. We will first prove a claim.

Claim 2.1. For I' an interval on alternate axis, we have

$$|I'|_e = |I'|'_e.$$

Proof. Take a slight blow up of I', denoted by I'^*, so that $I' \subseteq (I'^*)^{\circ}$ and

$$v(I'^*) \leqslant v(I') + \epsilon.$$

Since $(I'^*)^{\circ}$ is open, we may write

$$(I'^*)^{\circ} = \bigcup_k I_k$$

which are non-overlapping (see **Remark 1**). We then take a finite collection and note that we have

$$\sum_{k=1}^{N} v(I_k) \leqslant v(I'^*).$$

Thus, taking the limit as $N \to \infty$, we get

$$\sum_{k=1}^{\infty} v(I_k) \leqslant v(I'^*) \leqslant v(I') + \epsilon.$$

This tells us that $|I'|_e \leqslant v(I') + \epsilon$. Taking the limit as $\epsilon \to 0$, we get that

$$|I'|_e \leqslant v(I') = |I'|'_e.$$

An analogous argument gives us the reverse direction, and so we get equality.
Q.E.D

The above argument also gives us the following for free.

Corollary 2.2.1. For I an interval, we have

$$|I|_e = |I|'_e.$$

Take a general set $E \subseteq \mathbb{R}^n$ and cover it by "normal intervals" (that is, intervals on the standard axis); i.e. take $E \subseteq \bigcup_k I_k$ such that

$$\sum_k v(I_k) \leqslant |E|_e + \epsilon/2.$$

We can cover the I_k with alternate intervals. Take a collection $I'_{k,j}$ such that

$$I_k \subseteq \bigcup_j I'_{k,j},$$

$$\sum_j v(I'_{k,j}) \leqslant v(I_k) + (\epsilon/2)\, 2^{-k}.$$

By transitivity we get

$$E \subseteq \bigcup_{k,j} I'_{k,j}.$$

Moreover, we get

$$|E|'_e \leqslant \sum_{k,j} v(I'_{k,j}) \leqslant \sum_k \left(v(I_k) + (\epsilon/2)\, 2^{-k}\right) \leqslant |E|_e + \epsilon.$$

So, we get that $|E|'_e \leqslant |E|_e + \epsilon$ and moreover $|E|'_e \leqslant |E|_e$. A symmetric argument gives us the other direction. **Q.E.D**

We now are going to talk about **measurable sets**, or **Lebesgue measurable sets**. We will first, however, discuss a relation between outer measure and open sets.

Lemma 2.1. For any $E \subseteq \mathbb{R}^n$ and $\epsilon > 0$, there exists an open set G such that $E \subseteq G$ and

$$|G|_e \leqslant |E|_e + \epsilon.$$

Corollary 2.2.2. For any $E \subseteq \mathbb{R}^n$, we get that

$$|E|_e = \inf_{\substack{G \text{ open} \\ E \subseteq G}} |G|_e.$$

We now prove the lemma.

Proof. We what now seems to be the standard trick. Take a cover of $\{I_k\}$ of E such that

$$\sum_k v(I_k) \leqslant |E|_e + \epsilon/2.$$

Now take a blowup of the I_k such that $I_k \subseteq (I_k^*)^{\text{o}}$ and

$$v(I_k^*) \leqslant v(I_k) + \epsilon/2 2^{-k}.$$

Take G to be the union of the interiors of the blow up; that is,

$$G = \bigcup_k (I_k^*)^{\text{o}}.$$

This is open, and we have

$$|G|_e \leqslant \sum_k v\left((I_k^*)^{\text{o}}\right) \leqslant \sum_k \left(v(I_k) + (\epsilon/2)\, 2^{-k}\right) \leqslant |E|_e + \epsilon.$$

Q.E.D

This now leads us to the definition of a measurable set.

Definition. A set $E \subseteq \mathbb{R}^n$ is said to be **measurable** or **Lebesgue measurable** if for all $\epsilon > 0$ there is an open set G such that $E \subseteq G$ and $|G - E|_e < \epsilon$.

Definition. If E is measurable, then the **measure of E** is $|E|_e$; that is,

$$|E|_e = |E| = \mu(E).$$

Remark 6. Again, note that **Lemma 2.1** gave us **no** information on this definition. Just because $|G|_e - |E|_e < \epsilon$ **does not** imply that $|G - E|_e < \epsilon$. This highlights one of the issues with outer measure.

Remark 7. There are nonmeasurable sets, however most sets in your life are measurable (in other words, if you can write it down, it's measurable).

So what sets are exactly measurable, then?

Example 2.1. (i) **Open sets are measurable**. This is clear, taking $G = E$.

(ii) **Sets with outer measure 0 are measurable.**[1] Let $E \subseteq \mathbb{R}^n$ be such that $|E|_e = 0$. Then by **Lemma 2.1** we have that there is a G such that $E \subseteq G$ and $|G|_e \leqslant |E|_e + \epsilon$. Since $|E|_e = 0$, we get that $|G|_e \leqslant \epsilon$. Thus, we have

$$|G - E|_e \leqslant |G|_e \leqslant \epsilon.$$

So E is measurable.

(iii) **Countable unions of measurable sets are measurable**. Take $E = \bigcup_k E_k$. Take $\{G_k\}$ for each k such that $E_k \subseteq G_k$ and $|G_k - E_k|_e < \epsilon 2^{-k}$ (this is fine since the E_k are measurable). Let $G = \bigcup_k G_k$ Then we have

$$|G - E|_e \leqslant \left| \bigcup_k (G_k - E_k) \right|_e \leqslant \sum_k |G_k - E_k|_e < \epsilon.$$

(iv) **Intervals are measurable**. First, we prove a claim.

Claim 2.2. A set $E \subseteq \mathbb{R}^n$ consisting of a single point x has outer measure 0. Moreover, it is measurable.

Proof. Take the open sets

$$G_\epsilon = B_\epsilon(x) = \{y \in \mathbb{R}^n \ : \ |x - y| < \epsilon\}.$$

Then we have $E \subseteq G_\epsilon$ for all ϵ, and moreover

$$0 \leqslant |E|_e \leqslant \inf |G_\epsilon|_e = 0.$$

Q.E.D

Note that $I = \partial I \cup I^\circ$. We have I° open, and so measurable, and we have $\partial I = \bigcup\{x\}$ such that $x \in I - I^\circ$, which is a countable union of points. These have measure zero, and so ∂I is measurable and moreover has measure 0. Therefore, I is measurable. In particular, $|I| = |I^\circ|$.

[1] I slightly diverge from the lecture notes since I have an issue with the proof as I have it.

2.3 Lecture 3 (σ-algebra)

(v) **Closed sets are measurable**. First, we need a definition and a lemma.

Definition. Let $F, H \subseteq \mathbb{R}^n$ be two sets. Then we define the **distance** between the sets as

$$d(F, H) = \inf\{|x - y| \ : \ x \in F, y \in H\}.$$

Lemma 2.2. If $d(E_1, E_2) > 0$, then $|E_1 \cup E_2| = |E_1|_e + |E_2|_e$.

Proof. Let's say $E_1 \cup E_2 \subseteq \bigcup_k I_k$, where $\sum_k |I_k| < |E_1 \cup E_2|_e + \epsilon$, such that there is no $I_l \subseteq E_1 \cap E_2$ (do this by subdivisions). Thus, we have that $E_1 \subseteq \bigcup_l I_l$ and $E_2 \subseteq \bigcup_h I_h$. Then we get

$$|E_1|_e + |E_2|_e \leqslant \sum_l |I_l| + \sum_h |I_h| = \sum_k |I_k| < |E_1 \cup E_2|_e + \epsilon.$$

This implies that $|E_1|_e + |E_2|_e \leqslant |E_1 \cup E_2|_e$. We know the converse inequality is true by **Theorem 2.1 (iii)**. Thus, we have equality. **Q.E.D**

Lemma 2.3. If $\{I_k\}_{k=1}^N$ is a collection of nonoverlapping intervals, then

$$\left|\bigcup_{k=1}^N I_k\right| = \sum_{k=1}^N |I_k|.$$

Proof. We have by **Theorem 2.1 (iii)**

$$\left|\bigcup_{k=1}^N I_k\right| \leqslant \sum_{k=1}^N |I_k|,$$

so it suffices to show the other direction. Suppose that $\bigcup_{k=1}^N I_k$ is covered by intervals $\{J_j\}$. In other words,

$$\bigcup_{k=1}^N I_k \subseteq \bigcup_j J_j.$$

For $\epsilon > 0$ fixed and for each interval J_j, pick an interval J_j^* containing J_j in its interior, and such that

$$|J_j^*| \leqslant (1 + \epsilon)|J_j|.$$

Then since $\bigcup_{k=1}^N I_k$ is compact (it is bounded and a countable union of closed sets, and so closed) it is in fact covered by finitely many of the J_j^*. Hence, we have

$$\bigcup_{k=1}^N I_k \subseteq \bigcup_{j=1}^M J_j^*.$$

It follows then that

$$\left|\bigcup_{j=1}^{M} J_j^*\right| \geqslant \sum_{k=1}^{N} |I_k|.$$

On the other hand, we get

$$\left|\bigcup_{j=1}^{M} J_j^*\right| \leqslant \sum_{j=1}^{M} |J_j^*| \leqslant (1+\epsilon)\sum_{j=1}^{M} |J_j| \leqslant (1+\epsilon)\sum_{j} |J_j|.$$

Hence, for any cover $\{J_j\}$ of $\bigcup_{k=1}^{N} I_k$, we have

$$\frac{1}{1+\epsilon}\sum_{k=1}^{N} |I_k| \leqslant \sum_{j} |J_j|.$$

Take the infimum over all such covers and let $\epsilon \to 0$. **Q.E.D**

We will also establish two claims that are used implicitly in the next argument.

Claim 2.3. If G open, $F \subseteq G$ compact, then $G - F$ is open.

Proof. Notice that $G - F = G \cap F^c$, F compact implies it is closed so that F^c is open, and this is therefore the intersection of two open sets. **Q.E.D**

Claim 2.4. If F, H compact and $F \cap H = \varnothing$ (that is, they are disjoint), then $d(F, H) > 0$.

Proof. Assume otherwise. That is, assume $d(F, H) = 0$. Then per definition this says that

$$\inf\{|x - y| \ : \ x \in F, y \in H\} = 0.$$

But F and H being closed implies closed implies that there exists x, y such that $|x - y| = 0$, or $x = y$, since there exists $\{x_n\} \in F$, $\{y_n\} \in H$ such that $|x_n - y_n| \to 0$, but $x_n \to x \in F$ and $y_n \to y \in H$. Thus, we have $F \cap H \neq \varnothing$, a contradiction. **Q.E.D**

We now have enough tools to prove that closed sets are measurable. Throughout, let F be a closed set. We first start with the case where F is compact. Choose open G with $F \subseteq G$ and $|G| < |F|_e + \epsilon$, $\epsilon > 0$. Since $G - F$ is open, we can write it as

$$G - F = \bigcup_{k} I_k,$$

where the $\{I_k\}$ are nonoverlapping intervals. Thus, we get that

$$|G - F|_e \leqslant \sum_{k} |I_k|.$$

Note that

$$G = F \cup (G - F) = F \cup \left(\bigcup_k I_k \right).$$

We can then take a finite collection of the I_k, and note that

$$F \cup \left(\bigcup_{k=1}^N I_k \right) \subseteq G.$$

Notice as well that this is closed and bounded, and so therefore compact *and* disjoint. Hence, they have positive distance, and so we may use **Lemma 2.2** to get

$$|F|_e + \sum_{k=1}^N |I_k| \leqslant |G|.$$

Rewriting this, we get

$$\sum_{k=1}^N |I_k| \leqslant |G| - |F|_e < \epsilon.$$

Since this is true for all N, we may take the limit to get

$$|G - F|_e \leqslant \sum_k |I_k| \leqslant |G| - |F|_e < \epsilon.$$

For F not compact, write

$$F = \bigcup_k (F \cap B_k(x)).$$

Since this is a countable union of compact sets, we win.

(vi) **The complement of a measurable set is measurable**. For any $k \geqslant 1$, pick an open set G_k with $E \subseteq G_k$ and $|G_k - E|_e < 1/k$. If we look at the complement, we have $G_k^C \subseteq E^C$. Moreover, $\bigcup_k G_k^C \subseteq E^C$, and denote $\bigcup_k G_k^C = G$. Let $Z = E^C - H$. We claim $|Z|_e = 0$. Looking at Z, we see it is $Z = E^C \cap H^C = E^C \cap (\cap_k G_k)$. Certainly $Z \subseteq E^c \cap G_k = G_k - E^C$, and so we see $|Z|_e \leqslant |G_k - E|_e < 1/k$. Thus, we see $E^c = Z \cup H$, a union of two measurable sets.

(vii) The **Cantor set** is measurable. Take the interval $[0, 1]$, and subdivide it into thirds. Remove the interior of the middle third, leaving us with $\left(\frac{1}{3}, \frac{2}{3}\right)$. Each successive step follow this pattern; take each interval, subdivide it into thirds, and then remove the interior of the middle third. Denoting the set after the k^{th} step as C_k, we have the **Cantor set** C is what is left over after repeating this an infinite number of times; that is,

$$C = \bigcap_{k=1}^\infty C_k.$$

Since C_k is closed, it follows C is closed, and so (v) gives us that this is measurable. We see that $|C|_e = 0$, since

$$|C|_e \leqslant 2^k 3^{-k} \ \forall k.$$

Definition. A non-empty collection Σ of sets is called a **σ-algebra** if

(i) Σ is closed under countable unions.

(ii) Σ is closed under complements.

Remark 8. Notice that (i) and (ii) imply that it is closed under countable intersection as well.

If we let $\mathcal{M}$ be the collection of measurable sets, then $\mathcal{M}$ is a σ-algebra.

Definition. The **Borel σ-algebra** (denoted by $\mathcal{B}$) is the smallest σ-algebra containing the open sets.

Remark 9. The sets in $\mathcal{B}$ are measurable.

2.4 Lecture 4 (Limits)

We want to work with limits of sets as well. If $\{E_k\}_{k=1}^{\infty}$ is a sequence of sets such that $E_k \subseteq E_{k+1}$ for all k, then we define the limit of these sets to be the union. In other words, $E_k \nearrow \bigcup_k E_k$. The other direction is analogous with intersection and decreasing sets.

Definition. We define the $\liminf E_n$ to be

$$\liminf E_n = \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} E_n.$$

Analogously, we define the $\limsup E_n$ to be

$$\limsup E_n = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} E_n.$$

It's good to also notice what these actually mean in terms of words. We have that $\limsup E_n$ is the set of all elements such that the elements are in infinitely many E_n. Likewise, the $\liminf E_n$ is the set of all elements such that those limits are in E_n for all $n \geqslant n_0$, where n_0 can depend on that element.

Example 2.2. (i) Let $\{f_i\}$ be a sequence of continuous functions. Then we claim that

$$E = \{x \ : \ \lim_{n\to\infty} |f_n(x)| = 0\}$$

is measurable. To show this, we need to recall that the limit being zero implies that there is an $\epsilon > 0$ such that $|f_n(x)| < \epsilon$ for all $n \geqslant n_0$, where n_0 depends on x. Take $\epsilon = 1/k$. Then we may rewrite this all as

$$E = \bigcap_{k=1}^{\infty} \left\{ x \ : \ |f_n(x)| < \frac{1}{k} \ \forall n \geqslant n_0(x) \right\}.$$

However, the inside is a liminf! Using this, we may rewrite the whole thing as

$$E = \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \left\{ x \ : \ |f_n(x)| < \frac{1}{k} \right\}.$$

The set on the inside is open, since f is continuous, and so this is all just a bunch of unions and intersections of open sets. Since $\mathscr{M}$ is a σ-field, we get that this means E is measurable.

(ii) (Mandelbrot Set) Let $f_c(z) = z^2 + c$ over $\mathcal{C}$. Then the Mandelbrot set $M = \{c \in \mathcal{C} \ : \ f_c(0), f_c(f_c(0)), \ldots, f_c^{(n)}(0) \text{ is bounded sequence}\}$ is called the Mandelbrot set. We rewrite this as

$$M = \bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} \{c \in \mathbb{C} \ : \ |f_c^{(n)}(0)| \leqslant k\}.$$

The set on the inside is closed by an argument involving polynomial inequalities (also could note that it's the pullback on a closed set of a continuous function).

(iii) (Normal numbers) We have a **base-2 decimal expansion** for $x \in [0, 1)$ is defined to be

$$x = b_1 b_2 \ldots b_n \ldots = \sum_{n=1}^{\infty} \frac{b_i}{2^n}.$$

where $b_i \in \{0, 1\}$. Some numbers have two different base-2 decimal expansions, however we'll just consider the one ending in all 0s for simplicity. Let's look at

$$E = \{x \in [0, 1) \ : \ \text{decimal exp of x has equal number of 0s and 1s}\}.$$

Then we'd like to show that E is measurable. Let $r_n(x)$ be the number of 1's in the first n digits after the decimal point for the base-2 decimal expansion of x. Then

$$E = \left\{ x \in [0, 1) \ : \ \lim_{n \to \infty} \frac{1}{n} r_n(x) = \frac{1}{2} \right\}$$

$$= \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \left\{ x \in [0, 1) \ : \ \left| \frac{1}{n} r_n(x) - \frac{1}{2} \right| \leqslant \frac{1}{k} \right\}$$

$$= \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \left\{ x \in [0,1) \ : \ \frac{n}{2} - \frac{n}{k} \leqslant r_n(x) \leqslant \frac{n}{2} + \frac{n}{k} \right\}$$

$$= \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \bigcup_{r=\lceil \frac{n}{2} - \frac{n}{k} \rceil}^{\lfloor \frac{n}{2} + \frac{n}{k} \rfloor} E_{n,r},$$

where $E_{n,r}$ is the set of numbers in $[0,1)$ whose first n digits after the decimal point consist of exactly r ones and $n-r$ zeroes. We see that $E_{n,r}$ is measurable, since it is the disjoint union of $\binom{n}{r}$ intervals of the form $[j/2^n, (j+1)/2^n)$.

Now we would like to explore two properties of the Lebesgue measure.

Proposition 2.1. (i) If $\{E_k\}$ are disjoint measurable sets, then

$$\left| \bigcup_k E_k \right| = \sum_k |E_k|.$$

(ii) If $E_k \nearrow E$, then $\lim_{n\to\infty} |E_n| = |E|$. If $E_k \searrow E$ and at least one $|E_k| < \infty$, then $\lim_{n\to\infty} |E_n| = |E|$.

Before starting the proof, we need a lemma.

Lemma 2.4. We have that E is measurable if and only if for all $\epsilon > 0$ there exists an $F \subseteq E$ closed with $|E - F| < \epsilon$.

Proof. Since we are in a σ-field, we have that E is measurable if and only if E^c is measurable. Since E is measurable, we have that there is an open set G such that $E \subseteq G$ and $|G - E| < \epsilon$ for all $\epsilon > 0$. Let $G^c = F$. Then we have $F \subseteq E$ closed. Moreover,

$$G - E = G \cap E^c = F^c \cap E = E^c - F,$$

so that $|E^c - F| < \epsilon$. **Q.E.D**

We now prove the proposition.

Proof. (i) We first break this up into cases.

Case 1: Assume the E_k are bounded. Choose subsets $F_k \subseteq E_k$ closed with $|E_k - F_k| < \epsilon 2^{-k}$. Then we must have

$$|E_k| \leqslant |E_k - F_k| + |F_k| \leqslant |F_k| + \epsilon 2^{-k}. \tag{2.1}$$

Now, since the E_k are disjoint, we have that the F_k are disjoint. Moreover, since the E_k are bounded, then the F_k are also bounded. Since the F_k are closed, bounded, and disjoint, we get that they are compact and disjoint. Compact disjoint sets are separated (see **Claim 2.4**), and so we may use **Lemma 2.2** to get

$$\left| \bigcup_{n=1}^{N} F_n \right| = \sum_{n=1}^{N} |F_n|.$$

Noting that

$$\bigcup_k F_k \subseteq \bigcup_k E_k,$$

we get

$$\sum_{n=1}^{N} |F_n| \leqslant \left|\bigcup_k E_k\right|.$$

This is then true for arbitrary N, and so we can take the limit to get

$$\sum_k |F_k| \leqslant \left|\bigcup_k E_k\right|.$$

Using the inequality above, we then have

$$\sum_k \left(|E_k| - \epsilon 2^{-k}\right) = \sum_k |E_k| - \epsilon \leqslant \left|\bigcup_k E_k\right|.$$

The choice of ϵ is arbitrary, so letting it go to zero gives us the desired inequality.
Case 2: Now we do not assume the E_k are bounded. Let $\mathbb{R}^n = \bigcup_k I_k = \bigcup_k (I_k - I_{k-1})$, assuming $I_k \nearrow \mathbb{R}^n$. Now we take

$$E_k = \bigcup_j \left(E_k \cap (I_j - I_{j-1})\right).$$

This is disjoint and bounded, and so we may use the first case to get our desired inequality.

(ii) Assume $E_k \nearrow E$. Define $H_1 = E$, $H_k = E_k - E_{k-1}$ for $k \geqslant 2$. So $E = \bigsqcup_k H_k$. Note that $E_k = \bigsqcup_{j=1}^{k} H_j$. Because this is disjoint, we get

$$|E| = \sum_k |H_k| = \lim_{n\to\infty} \sum_{k=1}^{n} |H_k| = \lim_{n\to\infty} |E_n|.$$

Decreasing is a similar trick.

Q.E.D

2.5 Lecture 5 (Characterizations)

Definition. We say a set is of type G_δ if it is a countable intersection of open sets. Analogously, we say a set is of type F_σ if it is a countable union of closed sets.

Note that these are more general than open/closed sets. Notice as well that countable union/intersections preserve measure, and open/closed sets are measurable, so that sets of these types are measurable.

Theorem 2.3. The following statements are equivalent:

(i) E is measurable.

(ii) $E = H - Z$, where H is of type G_δ, $|Z| = 0$.

(iii) $E = H \cup Z$, where H is of type F_σ, $|Z| = 0$.

In other words, this characterization says that we can closely approximate measurable sets with G_δ and F_σ sets.

Proof. $(ii) \implies (i)$ and $(iii) \implies (i)$ are trivial; H and Z in both instances are measurable sets, and so differences of measurable sets and unions of measurable sets are measurable, thus giving us that E is measurable.
$(i) \implies (ii)$: Let G_k be a sequence of open sets such that $E \subseteq G_k$ and $|G_k - E| < 1/k$. Let $H = \bigcap_k G_k$. Clearly H is of type G_δ. Now, write $Z = H - E$ so that $E = H - Z$. So we are almost done; we just need to show that $|Z| = 0$. Notice that in particular we have that $Z \subseteq G_k - E$ for all k. Thus, we have $|Z| \leqslant |G_k - E| < 1/k$. Since this works for all k, we get that $|Z| = 0$.
$(i) \implies (iii)$: In this case, pick $F_k \subseteq E$ closed so that $|E - F_k| < 1/k$. Let $H = \bigcup_k F_k$ and write $Z = E - H$. Then this implies $E = Z \cup H$ and moreover we have $|Z| < |E - F_k|$ for all k, so in particular $|Z| < 1/k$ for all k. Thus, we have $|Z| = 0$, and we win. **Q.E.D**

Now, we want to discuss an alternative definition of measurability which will come up in later chapters called **Carathéodory's definition of measurability**.

Theorem 2.4. We have that E is measurable if and only if $|A|_e = |A \cap E|_e + |A - E|_e$ for all sets $A \subseteq \mathbb{R}^n$.

One thing to quickly note is that this does not rely on open or closed sets. As a result, this may be a useful characterization in spaces where the topology is not natural in any way. Another thing to note is that the A above does **not** need to be measurable in any way.

Proof. We first do the implication. Assume that E is measurable. Then by properties of outer measure we always have that $|A|_e \leqslant |A \cap E|_e + |A - E|_e$. We then want to show the other direction, that is, $|A \cap E|_e + |A - E|_e \leqslant |A|_e$. Start by choosing an open G such that $A \subseteq G$ and $|G| < |A|_e + \epsilon$. Then we have that $|G \cap E| + |A - E| = |G| < |E|_e + \epsilon$; moreover, by the inclusion relation, we have

$$|A \cap E|_e + |A - E|_e \leqslant |G \cap E| + |G - E| < |A|_e + \epsilon.$$

Since this is for any ϵ, we take it to go to zero and get $|A \cap E|_e + |A - E|_e \leqslant |A|_e$, as desired.

For the converse, we assume that $|A|_e = |A \cap E|_e + |A - E|_e$ for all subsets A. In particular, take $A = G$, an open subset such that $E \subseteq G$ and $|G| < |E|_e + \epsilon$. Then we have

$$|G \cap E|_e + |G - E|_e < |E|_e + \epsilon.$$

But we assume $E \subseteq G$, so we have $G \cap E = E$. Hence, we write this as

$$|E|_e + |G - E|_e < |E|_e + \epsilon \implies |G - E|_e < \epsilon.$$

Since this holds for arbitrary ϵ, we get that this is measurable by definition.
Q.E.D

The last thing we did was talk about the construction of a non-measurable set. I'll omit this, and just refer the reader to this. As a corollary to the construction from class, though, we get the following.

Corollary 2.4.1. If $|A|_e > 0$, then there exists a non-measurable subset of A.

Chapter 3

Functions and Measure on Euclidean Space

3.1 Lecture 6 (Functions that preserve measure)

We first recall what a Lipschitz function is.

Definition. We say that a function $T : \mathbb{R}^n \to \mathbb{R}^n$ is **Lipschitz** if

$$|T(x) - T(y)| \leqslant C|x - y|$$

for all x, y and $C \in \mathbb{R}$.

We also recall what the diameter of a set is.

Definition. The **diameter of a set** E is defined to be

$$\operatorname{diam}(E) = \sup\{|x - y| \ : \ x, y \in E\}.$$

This leads us to a theorem.

Theorem 3.1. If $T : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz and if $E \subseteq \mathbb{R}^n$ is measurable, then $T(E)$ is measurable. In short, Lipschitz functions preserve measurability.

We will first prove a lemma.

Lemma 3.1. If $f : X \to Y$ is a function, then for $\{U_\alpha\}$ such that $U_\alpha \subseteq X$ for all α we get

$$f\left(\bigcup_\alpha U_\alpha\right) = \bigcup_\alpha f(U_\alpha).$$

We will also need a claim for this lemma.

Claim 3.1. If f is continuous, E is compact, then $f(E)$ is compact (that is, the image of a compact set under a continuous function is compact).

Proof. Let $E \subseteq \bigcup_{n=1}^{M} f^{-1}(V_n)$. Then we have that

$$f(E) \subseteq \bigcup_{n=1}^{M} V_n.$$

Q.E.D

Proof. If $y \in f\left(\bigcup_\alpha U_\alpha\right)$, then we have $x \in \bigcup_\alpha U_\alpha$ such that $f(x) = y$. Thus, $x \in U_\alpha$ for some α, and we get $f(x) = y \in f(U_\alpha)$. But since this works for all α, we get

$$y \in \bigcup_\alpha f(U_\alpha).$$

The other directon analogous. **Q.E.D**

Remark 10. We have that

$$f\left(\bigcap_\alpha U_\alpha\right) \subseteq \bigcap_\alpha f(U_\alpha),$$

where equality may not hold.

Proof. We first show that T preserve F_σ sets. Suppose first we have a compact set. Since T is Lipschitz, it is also continuous, and so we have that $T(E)$ is also compact. If E is closed, we have that it is a union of compact sets, and since $T(\bigcup_k E_k) = \bigcup_k T(E_k)$, we have that $T(E)$ is also closed. Using this again, we get that if E is of type F_σ, then $T(E)$ is also of type F_σ.

Next, we quickly note that if $|Z| = 0$, then $|T(Z)| = 0$. Since T is Lipschitz, we have $\operatorname{diam}(T(E)) \leqslant C\operatorname{diam}(E)$. Then we have a constant C' such that $|T(I)|_e \leqslant C'|I|$. This gives us that $|T(E)|_e \leqslant C'|E|_e$. So if we have measure 0 on the right, we get measure 0 on the left.

This is all we need, since by **Theorem 2.3** if E is measurable we may write it as $E = H \cup Z$, where H is of type F_σ and $|Z| = 0$. Using **Lemma 3.1** again, we get that $T(E)$ is measurable. **Q.E.D**

So Lipschitz functions preserve measurable sets. The next theorem shows that we can do better than that.

Theorem 3.2. If T is a linear transformation of $\mathbb{R}^n$, then $|T(E)| = |\det(T)| \cdot |E|$ for all measurable $E \subseteq \mathbb{R}^n$.

Proof. Recall from linear algebra that

$$|T(I)| = |\det(T)| \cdot |I|.$$

From this, it follows that $|T(E)|_e \leqslant |\det(T)| \cdot |E|_e$. So for measure 0 sets the conclusion follows. Without loss of generality, we may assume that $\det(T) > 0$ since the conclusion also follows from this. Since $\det(T) > 0$, we have that T is invertible. Now, note that theorem is true for open sets, since for open U we may write it as

$$U = \bigcup_k I_k,$$

where the I_k are non-overlapping. Hence, we have

$$T(U) = \bigcup_k T(I_k),$$

and so

$$|T(U)| = \det(T) \sum_k |T(I_k)|.$$

Let's now explore sets of type G_δ. Let $H = \bigcap_k G_k$, and let $H_n = \bigcap_{k=1}^n G_k$. Then $H_n \searrow H$. Assuming $|H_n| \to |H|$ (that is, at least one of the H_i has finite measure), we know $T(H_n) \searrow T(H)$, using the fact that T is invertible. So we know that $|T(H_n)| \to |T(H)|$, and we know that $T(H_n) = \det(T) \cdot |H_n|$, but this converges to $T(H) = \det(T) \cdot |H|$.

Why may we assume that $|H_n| \to |H|$? Write

$$H = \bigcup_{n=1}^{\infty} (H \cap \{x \in \mathbb{R}^n \ : \ |x| < n\}) = \bigcup_{n=1}^{\infty} E_n.$$

Each of the E_n is of type G_δ, where the above argument works. We have $E_n \nearrow H$, and so $|E_n| \nearrow |H|$.

Finally, let E be an arbitrary measurable set. Again, by **Theorem 2.3**, we have that we may write $E = H \backslash Z$, where H is of type G_δ and $|Z| = 0$. We also take $E \subseteq H$. Then we have $T(E) = T(H) - T(Z)$, using the fact that T is invertible. So

$$|T(E)| = |T(H)| - |T(Z)| = |T(H)| = \det(T) \cdot |H|.$$

Note that $|H| = |E|$ by assumption, so we get $|T(E)| = \det(T) \cdot |E|$. **Q.E.D**

Remark 11. The **Cantor Function** is related to the Cantor set, $C = \bigcap_n C_n$. Let $D_n = [0,1] - C_n = \bigcup_{j=1}^{2^n-1} \left(I_j^n\right)^{\mathrm{o}}$. Define a function

$$f_n(x) = \begin{cases} j/2^n \text{ if } x \in I_j^n \\ \text{linear otherwise} \end{cases}.$$

This function is continuous, but maps a measurable set to a non-measurable set. To see this, let $f(x) = \lim_{n\to\infty} f_n(x)$. Notice that f surjects onto $[0,1]$, so use Vitali's theorem to find a non-measurable set $V \subseteq [0,1]$. Then we have that $f^{-1}(V) \subseteq C$, and so $|f^{-1}(V)| \leqslant |C| = 0$. Therefore, $f^{-1}(V)$ is measurable, and f maps this set to V, a non-measurable set.

3.2 Lecture 7 (Measurable Functions)

The book defines a measurable function as follows.

Definition. Given a subset $E \subseteq \mathbb{R}^n$ (not necessarily measurable) and a function $f : E \to \mathbb{R}$, we say that f is a **measurable function** on E if $\{x \in E \ : \ f(x) > a\}$ is a measurable set for all $a \in \mathbb{R}$.

We will use a similar definition, given below.

Definition. A function $f : \mathbb{R}^n \to \mathbb{R}$ is a **measurable function** if $\{x \ : \ f(x) > a\}$ is measurable for all $a \in \mathbb{R}$.

We abbreviate the pullback $\{x \ : \ f(x) > a\}$ to simply be $\{f > a\}$ throughout.

Example 3.1. (i) We have that **continuous functions** are measurable functions. The set (a, ∞) is an open set, and so we have $f^{-1}((a, \infty)) = \{f > a\}$ is an open set, and so by **Example 2.1 (i)** we have that it is measurable.

(ii) The **characteristic function of measurable sets** is measurable. Recall that for a set A we define the characteristic function as

$$\chi_A(x) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ if } x \notin A \end{cases}.$$

Then if $E \subseteq \mathbb{R}^n$ is measurable, we have χ_E is a measurable function. First, take $a \geqslant 1$. We have then that $\chi_E^{-1} = \varnothing$, and so measurable. For $0 \leqslant a < 1$, we have $\chi_E^{-1} = E$, a measurable set. Finally, if $a \leqslant 0$, we have that $\chi_E^{-1} = \mathbb{R}^n$, a measurable set.

The choice of $\{f > a\}$ is not unique, as we see in the following theorem.

Theorem 3.3. The following are conditions are equivalent for a function f.

(i) $\{f > a\}$ is measurable for all $a \in \mathbb{R}$.

(ii) $\{f \geqslant a\}$ is measurable for all $a \in \mathbb{R}$.

(iii) $\{f < a\}$ is measurable for all $a \in \mathbb{R}$.

(iv) $\{f \leqslant a\}$ is measurable for all $a \in \mathbb{R}$.

Proof. $(i) \implies (ii)$: We have that

$$\{f \geqslant a\} = \bigcap_{k=1}^{\infty} \{f > a - 1/k\}.$$

Since this is a countable intersection of measurable sets, it is also measurable. $(ii) \implies (iii)$: Measurability is preserved under complements, and we have that

$$\{f \geqslant a\}^C = \{f < a\}.$$

$(iii) \implies (iv)$: Again, we see that

$$\{f \leqslant a\} = \bigcup_{k=1}^{\infty} \{f < a + 1/k\}.$$

A countable union of measurable sets is measurable.
$(iv) \implies (i)$: Measurability is preserved under complements, and we have that

$$\{f \leqslant a\}^C = \{f > a\}.$$

Q.E.D

Using this, we can construct some nicer measurable sets.

Corollary 3.3.1. If f a measurable function, then the following are measurable sets:

(i) $\{a \leqslant f \leqslant b\}$,

(ii) $\{f = a\}$,

(iii) $\{a < f \leqslant b\}$ and $\{a \leqslant f < b\}$,

(iv) $\{f < \infty\}$.

We can also obtain some alternative criteria for measurability using this definition.

Theorem 3.4. We have that f is a measurable function if and only if $f^{-1}(G)$ is a measurable set for all $G \subseteq \mathbb{R}$ open.

Proof. $\impliedby$: is clear. We have that (a, ∞) is an open set for all $a \in \mathbb{R}$, and so furthermore $f^{-1}((a, \infty))$ is a measurable set for all a. By definition, this means that f is a measurable function.
$\implies$: Let $G \subseteq \mathbb{R}$ be open. Then since we are in $\mathbb{R}$, we may use **Remark 1** to get that there is an open cover of G of open disjoint intervals. Notationally, we have

$$G = \bigsqcup_k (a_k, b_k).$$

Now, the pullback of G is

$$f^{-1}(G) = \bigsqcup_k f^{-1}((a_k, b_k)).$$

Notice that we preserve the disjointness since we are in $\mathbb{R}$. The pullback of these open intervals is measurable since f is a measurable function, and so we have a disjoint union of measurable sets, which is measurable. **Q.E.D**

We also note that we need not check the condition for all $a \in \mathbb{R}$, but rather just for $a \in A \subseteq \mathbb{R}$ where $\mathrm{Cl}(A) = \mathbb{R}$ (i.e. a dense subset).

Theorem 3.5. If A is a dense subset of $\mathbb{R}$, then f is a measurable function if and only if $\{f > a\}$ is measurable for all $a \in A$.

Proof. $\implies$: is clear. By definition, we have $\{f > a\}$ is measurable for all $a \in \mathbb{R}$, and so by extension for all $a \in A$.
$\impliedby$: What if $a \notin A$? Since A is dense, we can take a sequence $\{a_k\}$ such that $a_k \searrow a$. Then we have

$$\{f > a_k\} \nearrow \{f > a\}.$$

So we have a limit of increasing measurable sets, and so by **Proposition 2.1 (ii)** we get that $\{f > a\}$ is measurable. **Q.E.D**

Definition. We say that a property holds **almost everywhere** (abbreviated by a.e.) if it holds everywhere except in some set of measure 0.

Example 3.2. (i) Let $f(x) = \chi_{\mathbb{Q}}(x)$. We claim $f = 0$ a.e. We see it holds everywhere except on $\mathbb{Q}$. So we must show $\mathbb{Q}$ has measure 0.

Claim 3.2. The measure of $\mathbb{Q}$ is 0.

Proof. We may write

$$\mathbb{Q} = \bigsqcup_{q \in \mathbb{Q}} \{q\}.$$

Since $\mathbb{Q}$ is countable, we have that this is a countable union of sets of measure 0. Furthermore,

$$|\mathbb{Q}| = \sum_{q \in \mathbb{Q}} |\{q\}| = \sum_{q \in \mathbb{Q}} 0 = 0.$$

Q.E.D

Thus, we have the property holds a.e.

(ii) We say $f = g$ a.e. if $|\{x \ : \ f(x) \neq g(x)\}| = 0$.

We may extend **Example 3.2 (ii)** to show that if one of the functions is measurable, then the other must also be measurable.

Theorem 3.6. If f is a measurable function, and $f = g$ a.e., then g is measurable and $|\{f > a\}| = |\{g > a\}|$ for all a.

Proof. Let $Z = \{x \ : \ f(x) \neq g(x)\}$. Then since $f = g$ a.e., we have $|Z| = 0$. Furthermore, for all a, we have

$$\{g > a\} \cup Z = \{f > a\} \cup Z.$$

Since $\{g > a\} \cup Z$ is measurable and differs from $\{g > a\}$ by a set of measure 0, we have that $\{g > a\}$ is measurable and shares the same measure. Furthermore,

$$|\{g > a\}| = |\{g > a\} \cup Z| = |\{f > a\} \cup Z| = |\{f > a\}|.$$

Q.E.D

Do these measurability conditions hold under composition? The answer is sometimes.

Theorem 3.7. If $\phi : \mathbb{R} \to \mathbb{R}$ is continuous and $f : \mathbb{R}^n \to \mathbb{R}$ is measurable, then $\phi \circ f$ is measurable.

Proof. We have the following diagram:

$$\begin{array}{ccccc} \mathbb{R}^n & \xrightarrow{\quad f \quad} & \mathbb{R} & \xrightarrow{\quad \phi \quad} & \mathbb{R} \\ \uparrow & & \uparrow & & \uparrow \\ f^{-1}(\phi^{-1}(G)) = G'' & \xleftarrow[f^{-1}]{} & \phi^{-1}(G) = G' & \xleftarrow[\phi^{-1}]{} & G \end{array}$$

Notice that G' is open, since ϕ is continuous, and G'' is measurable, since f is a measurable function. Thus, we have that $\phi \circ f$ is measurable. **Q.E.D**

Corollary 3.7.1. If f is measurable, then so are the following:

(i) $|f|$,

(ii) $|f|^p$,

(iii) e^{cf}, where c is a constant,

(iv) $f^+ = \max\{0, f\}$,

(v) $f^- = -\min\{0, f\}$.

Remark 12. Note that **Theorem 3.7** is **not** true in general. We are not guaranteed $\phi \circ f$ is measurable if ϕ is measurable, since we only know that $\phi^{-1}(G)$ is a measurable set, not open.

3.3 Lecture 8 (Egorov's Theorem)

We now prove some properties about measurable functions.

Theorem 3.8. (i) If f is a measurable function, $c \in \mathbb{R}$, then cf is measurable.

(ii) If f and g are measurable functions, then so is $f + g$.

(iii) If f is a measurable function, then so is f^2.

(iv) If f and g are measurable functions, then so is fg.

(v) If f and g are measurable functions, and $g \neq 0$ a.e., then f/g is measurable as well.

(vi) If $\{f_n\}$ is a sequence of measurable functions, then $\inf f_n$ and $\sup f_n$ are also measurable.

(vii) If $\{f_n\}$ is a sequence of measurable functions, then $\limsup$ and $\liminf$ are also measurable.

(viii) If $\lim_{n\to\infty} f_n = f(x)$ a.e., then f is measurable.

Proof. (i) We would like to show that the set $\{cf > a\}$ is measurable for all $a \in \mathbb{R}$. We may rewrite this as $\{f > a/c\}$, and since f is measurable we know that this set is measurable as well for all $a \in \mathbb{R}$.

(ii) We first need a lemma (most likely trivial and should be known, but I'll leave it in for completeness):

Lemma 3.2. Between any two real numbers is a rational number.

Proof. Assume without loss of generality that $x, y > 0$ (the argument is analogous in the other cases). Let $x, y \in \mathbb{R}$ such that $y > x$. Then we have $y - x > 0$. By the Archimedean Principle, we have that there exists an $n \in \mathbb{N}$ such that $y - x > 1/n > 0$. Choose the largest $k \in \mathbb{N}$ so that

$$\frac{k}{n} \leqslant x.$$

Then since this was the largest, we have

$$x < \frac{k+1}{n}.$$

Assume that $y \leqslant \frac{k+1}{n}$. Then we have

$$y - x \leqslant \frac{k+1}{n} - \frac{k}{n} = \frac{1}{n},$$

which is a contradiction. Hence, we must have that $y > \frac{k+1}{n}$, and therefore $x < \frac{k+1}{n} < y$ strictly. **Q.E.D**

We would like to show $\{f + g > a\}$ is measurable for all a. Rewrite this as $\{f > a - g\}$. Since $a - g$ and f are real numbers for all x, we get that there is at least one rational q in between them. In particular, we may rewrite this as

$$\bigcup_{q\in\mathbb{Q}} (\{f > q\} \cap \{q > a - g\}) = \bigcup_{q\in\mathbb{Q}} (\{f > q\} \cap \{g > a - q\}).$$

We know that these two sets are measurable, and so therefore this is a union of measurable sets and so measurable for all a.

(iii) We would like to show that $\{f^2 > a\}$ is measurable. We may rewrite this as $\{f^2 > a\} = \{f > \sqrt{a}\} \cup \{f < -\sqrt{a}\}$. We know that these sets are measurable, and so the original set is measurable.

(iv) Notice that $(f+g)^2 - (f-g)^2 = f^2 + 2fg + g^2 - f^2 + 2fg - g^2 = 4fg$. So we have
$$fg = \frac{(f+g)^2 - (f-g)^2}{4}.$$
We know that measurability is preserved under addition, subtraction, squares, and multiplication by constants. Hence, fg is measurable.

(v) If g is a measurable function such that $g \neq 0$ a.e., then we may define a new function h, which is equal to g a.e. and is equal to 1 where $g = 0$. Since $h = g$ a.e., we get that h is measurable. Furthermore, we see that we clearly get that f/h is measurable, and since $f/g = f/h$ a.e. we have that f/g is measurable by **Theorem 3.6**.

(vi) Notice that
$$\left\{\sup_k f_k > a\right\} = \bigcup_k \{f_k > a\},$$
$$\left\{\inf_k f_k < a\right\} = \bigcap_k \{f_k < a\},$$
and so we get that these functions are measurable.

(vii) Note that $\limsup f_n(x) = \inf_m \sup_{n \geqslant m} f_n(x)$ and $\liminf f_n(x) = \sup_m \inf_{n \geqslant m} f_n(x)$. Then by the prior property we get that these are measurable.

(viii) If $\lim_{n\to\infty} f_n(x) = f(x)$ a.e., then we get that it is equal to the $\limsup$ a.e., and so in particular we get that this f is measurable.

Q.E.D

Using these properties, we can build some crazy things. Moreover, we will show that even though we can get some crazy functions, these functions are well approximated by some nice functions.

Definition. A **simple function** is a function taking on only finitely many finite values. In other words, it is a function taking on $a_1, \ldots, a_N \in \mathbb{R}$ on sets $E_1, \ldots, E_N$ which are disjoint and whose union is $\mathbb{R}^n$. We may represent this function then by
$$f = \sum_{i=1}^{N} a_i \chi_{E_i}(x).$$

Lemma 3.3. We have that a simple function f is measurable if and only if the E_i are measurable for all i.

Proof. This is a relatively clear lemma. If f is measurable, then we have that $\{x : f(x) > a\}$ is measurable for all a. But this can be represented as a union of the E_i, and so we must have that the E_i are all measurable. For the other direction, if the E_i are all measurable, then we have $\{x : f(x) > a\} = \bigsqcup_{i=1}^{k} E_i$ for all a, where k depends on a. This is going to then be measurable for all a. **Q.E.D**

We now show that these simple functions are good approximators of all functions. We will use a sort of "Lebesgue philosophy" to do so. The following theorem will be extremely important moving forward.

Theorem 3.9. Every measurable function f can be expressed as the limit of simple measurable functions; i.e.

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

Moreover, if the $f \geqslant 0$ for all x, then we can choose the f_n so that $f_n \nearrow f$.

Proof. We first assume $f \geqslant 0$. Let

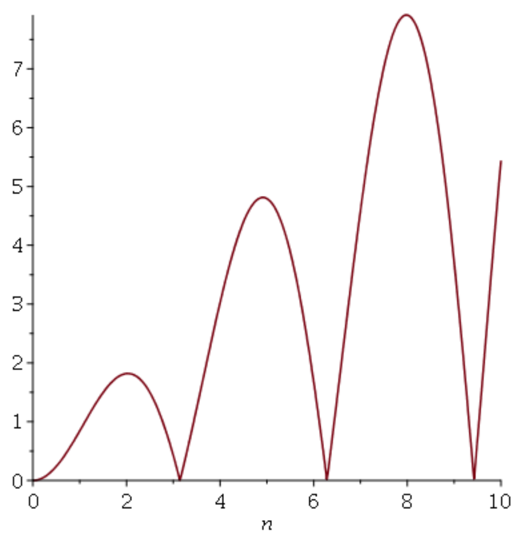
$$f_n(x) = \sum_{j=1}^{n2^n} \left(\left(\frac{j-1}{2^n} \right) \chi_{\{\frac{j-1}{2^n} < f \leqslant \frac{j}{2^n}\}} \right) + n\chi_{\{f \geqslant n\}}.$$

What is this function doing? It divides up the plane horizontally into dyadic intervals of length 2^n from 0 to n. Within each of these intervals, it takes the smallest value of f within this interval. If the function then goes above n, it just cuts off the function there. Certainly, then, we see that $f_n \nearrow f$, and we achieve our desired result. What if, however, $f < 0$ at some points? We may write $f = f^+ - f^-$ (see **Corollary 3.7.1**). Both of these functions are non-negative, and so we just use the prior argument to get the desired result. One may worry that we might get $\infty - \infty$; however notice that if one function is non-zero at some point, the other must be zero. **Q.E.D**

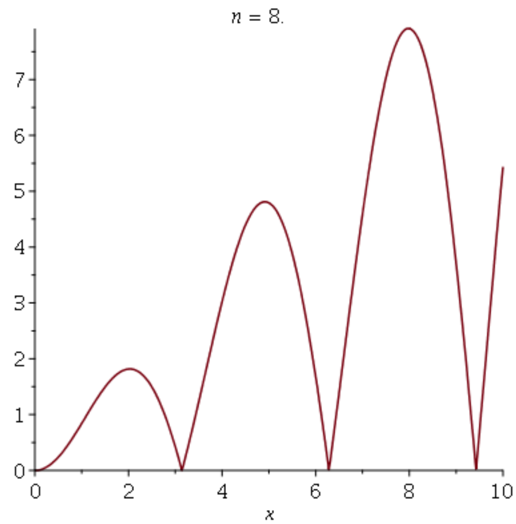
Example 3.3. To see this convergence, let's animate an example. Let

$$f(x) = |x \sin(x)|.$$

Using Maple, we plot this graph;



Animating from $0 \leqslant x \leqslant 10$ and $1 \leqslant n \leqslant 8$, we get the following gif. The end results is as follows;



Thus, it seems pretty evident that the sequence of functions converge.

3.4 Lecture 9 (Semi-Continuity)

Definition. Suppose that we have $f : E \to \mathbb{R}$. Then we say f is **upper semi-continuous** (abbreviated by usc) at $x_0 \in E$ if

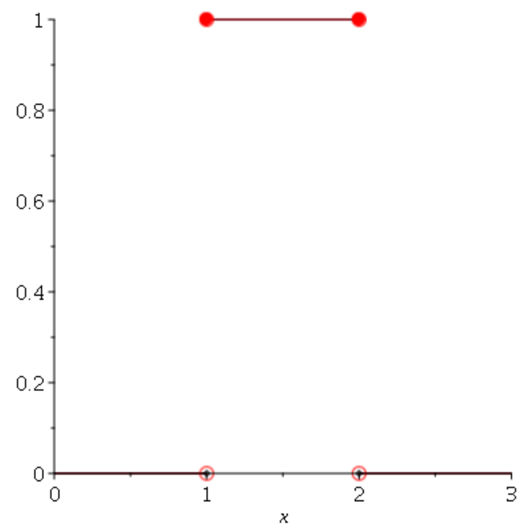
$$\lim_{x \to x_0, x \in E} \sup f(x) \leqslant f(x_0).$$

Likewise, we say that f is **lower semi-continuous** (abbreviated lsc) at $x_0 \in E$ if

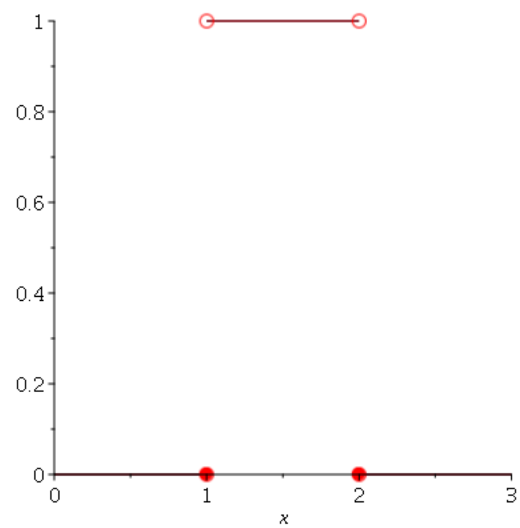
$$\lim_{x \to x_0, x \in E} \inf f(x) \geqslant f(x_0).$$

In other words, we have that f is usc if whenver there is a jump, we take the uppermost value, and it is lsc if whenever there is a jump we take the lowermost value.

Example 3.4. (a) Suppose $f = \chi_F$, $F \subseteq \mathbb{R}^n$ closed. Then f is usc. To see this, say $x_0 \notin F$. If $x_n \to x_0$ is some sequence of points converging to x_0, then we must have $x_n \in F^c$ by the closed property for all n large. So then $\chi_F(x_n) = 0$ for all n large enough; hence, $\limsup_n \chi_F(x_n) = 0 = \chi_F(x)$. If $x_0 \in F$, then we have that $\chi_F(x_0) = 1$. Since this is the characteristic function, we trivially get that $\chi_F(x_0) \geqslant \limsup_x \chi_F(x)$, since we have that the rightmost value can be at most 1. Below is a picture of this kind of a case for $\chi_{[1,2]}$



(b) Analogously, if $f = \chi_G$, $G \subseteq \mathbb{R}^n$ open, then f is lsc. Below is a picture of this kind of case for $\chi_{(1,2)}$.



We now discuss an alternate characterization of usc and lsc.

Theorem 3.10. (i) We have that a function f is usc if and only if the sets $\{f \geqslant a\}$ are closed for all $a \in \mathbb{R}$.

(ii) We have that a function f is lsc if and only if $\{f \leqslant a\}$ are closed for all $a \in \mathbb{R}$.

Proof. We first note that (i) and (ii) are equivalent, since the negative of a usc is a lsc function and vice versa.
($\Longrightarrow$) Assume that f is usc. Let $a \in \mathbb{R}$ be fixed, and let $x_0 \in \overline{\{f \geqslant a\}}$; i.e. the closure. Then we want to show $x_0 \in \{f \geqslant a\}$. Since x_0 is in the closure, we have a sequence of $x_n \in \{f \geqslant a\}$ which converge to x_0. Then we have $f(x_n) \geqslant a$ for all n. Since f is usc, then we have

$$\limsup_{n\to\infty} f(x_n) \leqslant f(x_0).$$

Therefore, we can chain this together to get

$$a \leqslant \limsup_{n\to\infty} f(x_n) \leqslant f(x_0).$$

But this implies that $f(x_0) \geqslant a$, which gives us that $x_0 \in \{f \geqslant a\}$. So we have that the set is equal to it's closure, and so is closed.
($\Longleftarrow$) Let x_0 be a limit point of E that is in E. If f is not usc at x_0, then $f(x_0) < \infty$ and there exists M and $\{x_k\}$ such that $f(x_0) < M$, $x_k \in E$, $x_k \to x_0$, and $f(x_k) \geqslant M$. Hence, $\{x : f(x) \geqslant M\}$ is not relatively closed since it does not contain all its limit points that are in E. So by contradiction we get that $f(x_0)$ is usc. **Q.E.D**

We now want to talk about Egorov's Theorem. The idea behind Egorov's Theorem is that almost everywhere convergence implies uniform convergence on a "large" subset. We give the more formal statement below.

Theorem 3.11. (Egorov) If $f_n \to f$ a.e. on E and $|E| < \infty$, then for all $\epsilon > 0$ there exists $F \subseteq E$ closed with $|E \backslash F| < \epsilon$ such that $f_n \to f$ uniformly on F.

Example 3.5. Take the function $f_n(x) = x^n$. Then we see that it converges to 0 uniformly on $(-1, 1)$. So we have that for all $\delta > 0$, $f_n \to 0$ uniformly on $[-1 + \delta, 1 - \delta]$.

Before proving the theorem, we need a lemma.

Lemma 3.4. Suppose that $\{f_n\}$ is a sequence of measurable functions that converges almost everywhere on a set E, $|E| < \infty$, to a function f. Then given $\epsilon, \eta > 0$, there is a closed subset F of E and an integer K such that $|E \backslash F| < \eta$ and $|f(x) - f_k(x)| < \epsilon|$ for all $x \in F$ and $k > K$.

Proof. Fix $\epsilon, \eta > 0$. For each m, let $E_m = \{|f_k - f| < \epsilon \ \forall k > m\}$. Thus,

$$E_m = \bigcap_{k>m} \{|f_k - f| < \epsilon\},$$

so that E_m is measurable. Clearly, we get that $E_m \subset E_{m+1}$. Moreover, since $f_k \to f$ a.e. in E and f is finite, $E_m \nearrow E \backslash Z$, where $|Z| = 0$. Hence, we have that $|E_m| \to |E \backslash Z| = |E|$. Since $|E| < \infty$, we have that $|E - E_m| \to 0$. We may thus choose an m_0 so that $|E - E_{m_0}| < \frac{1}{2}\eta$, and let F be a closed subset of E_{m_0} with $|E_{m_0} - F| < \frac{1}{2}\eta$. Then $|E - F| < \eta$, and $|f - f_k| < \epsilon$ in F if $k > m_0$. **Q.E.D**

The lemma is almost the same thing as the theorem, except we need to show uniform convergence. We now proceed to use the lemma to prove this.

Proof. Given $\epsilon > 0$, for each $k \geqslant 1$, choose closed $F_k \subseteq E$ and $m \geqslant 1$, such that

(i) $|E \backslash F_k| < \epsilon 2^{-k}$,

(ii) $F_k \subseteq \bigcap_{n=m}^{\infty} \{|f_n - f| < 1/k\}$.

By the lemma, we know that we may do this. Let $F = \bigcap_{k=1}^{\infty} F_k$. Then F is a closed subset of E, and we need to check that the statements of the theorem hold. First, we see that $|E\backslash F| < \epsilon$, since $E \backslash \bigcap_{k=1}^{\infty} F_k = \bigcup_{k=1}^{\infty} |E \backslash F_k|$. Next, we need to check the uniform convergence property. Given $\delta > 0$, pick $k \geqslant 1$ such that $1/k \leqslant \delta$. Then $F \subseteq F_k$. If $x \in F$, then $|f_n(x) - f(x)| < 1/k$ for all $n \geqslant m$. So $|f_n - f| < \delta$ by chaining these inequalities together. Thus, we get uniform convergence. **Q.E.D**

3.5 Lecture 10 (Lusin's Theorem)

We now want to go into Lusin's theorem. Intuitively, we have that Lusin's theorem says that "measurable functions are *almost* continuous." To properly say this, though, we need a few definitions.

Definition. We say that a function $f : X \to Y$ is **continuous relative to a set E** if for all $x_0 \in E$, we have

$$\lim_{x \to x_0, x \in E} f(x) = f(x_0).$$

Definition. A function f has **property $\mathcal{C}$ on E** if, given $\epsilon > 0$, there exists a closed subset $F \subseteq E$ with

(i) $|E \backslash F| < \epsilon$.

(ii) f is continuous relative to F.

This leads us to the following lemma.

Lemma 3.5. Simple measurable functions have property $\mathcal{C}$.

Proof. Let $\{E_i\}_{i=1}^{N}$ be a collection of disjoint sets, and

$$f = \sum_{i=1}^{N} a_i \chi_{E_i}.$$

Let $E = \bigsqcup_{i=1}^{N} E_i$ as well. For each i, choose closed subsets $F_i \subseteq E_i$ such that

$$|E_i \backslash F_i| < \epsilon/N.$$

Let $F = \bigsqcup_{i=1}^N F_i$. Then we have

$$|E\backslash F| = \left|\bigcup_{i=1}^N (E_i\backslash F_i)\right| \leqslant \sum_{i=1}^N |E_i\backslash F_i| < \sum_{i=1}^N \frac{\epsilon}{N} = \epsilon.$$

Thus, we just need to check that f is relatively continuous on F. Take $x_0 \in F$. Then we want to show

$$\lim_{x\to x_0, x\in F} f(x) = f(x_0).$$

Since the F_i are disjoint, we have $x_0 \in F_i$. Thus, since $x_n \to x_0$ and F_i closed, we have $x_n \in F_i$ for n large enough. Thus, we must get

$$\lim_{x\to x_0, x\in F} f(x) = a_i = f(x_0).$$

This works for all $x_0 \in F$ and so we get that the statement holds. **Q.E.D**

We now have enough to prove and state Lusin's theorem.

Theorem 3.12. (Lusin) If f is a finite function on measurable E, then f is measurable if and only if f has property $\mathcal{C}$.

Proof. ($\Longleftarrow$) Assume f has property $\mathcal{C}$. Then for all $k \geqslant 1$, we can choose $F_k \subseteq E$ so that

$$|E\backslash F_k| < 1/k$$

and f is relatively continuous on F_k. Let $H_n = \bigcup_{k=1}^n F_k$. Then H_n is closed, since we have a countable union of closed sets, and $H_n \subseteq H_{n+1}$. Notice as well that

$$|E\backslash H_n| < 1/n,$$

since $F_n \subseteq H_n$ and

$$|E\backslash H_n| < |E\backslash F_n| < 1/n.$$

Notice as well that f is relatively continuous on H_n. Then we see that $H_n \nearrow H \subseteq E$ with $|E\backslash H| = 0$. Now, let

$$g_n(x) = \begin{cases} f(x) \text{ if } x \in H_n, \\ 0 \text{ otherwise.} \end{cases}$$

Then $\lim_{n\to\infty} g_n(x) = f(x)$ for all $x \in H$. Notice that g_n is measurable as well; examining $\{g_n > a\}$, if we have $a > 0$ then we have it pulls back to $H_n \cap G$ by relative continuity, and for $a \leqslant 0$ we have that it pulls back to everything. Notice as well that since $|E\backslash H| = 0$, we get that $\lim_{n\to\infty} g_n(x) = f(x)$ a.e. By **Theorem 3.8 (vii)**, we see that f is then measurable.
($\Longrightarrow$) We assume that f is measurable. By **Theorem 3.9**, we may write it as the limit of a sequence of $\{f_n\}$ where f_n is a simple measurable function. We then break this up into cases.

Case 1: Assume $|E| < \infty$. By **Lemma 3.5**, we see that each f_n has property $\mathcal{C}$, and so we can pick $F_n \subseteq E$ closed such that

$$|E \backslash F_n| < \frac{\epsilon}{2} 2^{-n}$$

and f_n is relatively continuous on F_n. We need to have an additional set as well; take $F_0 \subseteq E$ closed such that

$$|E \backslash F_0| < \frac{\epsilon}{2},$$

and $f_n \to f$ uniformly on F_0. We may do so by Egorov's theorem, **Theorem 3.11**. Let us take then $F = \bigcap_{k=0}^{\infty} F_k \subseteq E$. Then F is closed, and $|E \backslash F| < \epsilon$. We need to show that f is continuous on F. On F, we have that the f_n are all relatively continuous, and $f_n \to f$ uniformly. Since this is uniform, we have that f is continuous on F.
Case 2: We assume now that $|E| = \infty$. Write $E = \bigcup E_k$, where

$$E_k = E \bigcap \{x \ : \ k-1 < |x| < k\};$$

i.e., E intersected with the set of washers. Choose closed $F_k \subseteq E_K$ where $|E_k \backslash F_k| < \epsilon 2^{-k}$, and f continuous on F_k. Now, let $F = \bigcup_{k=1}^{\infty} F_k$. We almost win; we just need to check that F is closed, and that f is continuous on F.

Let's first show that F is closed. Examine $x \in \bar{F}$, and take a sequence $x_n \to x$ such that $x_n \in F$. We see that x_n must eventually be in some washer F_k for some k and n large. Therefore, we get that $x \in F_k$, and so $x \in F$. Continuity follows easily from this. **Q.E.D**

Remark 13. We used the standard trick of getting infinite sets from unions of smaller sets which we understand well, but notice that we needed to be clever with the choice of our smaller sets.

We now want to discuss convergence in measure, which is weaker form of convergence.

Definition. We say that a sequence of functions $\{f_n\}$ **converges in measure on E** to a function f, written $f_n \xrightarrow{m} f$, if for all $\epsilon > 0$ we have

$$|\{x \in E \ : \ |f_n(x) - f(x)| > \epsilon\}| \to 0.$$

Remark 14. The reason this is weaker is that this doesn't necessarily imply the same thing for fixed points. It just works for overall collections.

Example 3.6. Let Q_k be the enumeration of the rationals in $[0,1]$. Let $f_k = \chi_{[q_k, q_k + 1/k]}$. Then $f_k \xrightarrow{m} 0$ only.

The next theorem says that this sequence gives us almost everywhere convergence for a subsequence.

Theorem 3.13. If $f_n \xrightarrow{m} f$ on E, then there exists a subsequence $\{n_k\}$ with $f_{n_k} \to f$ almost everywhere.

Proof. For each $k \geqslant 1$, choose $\{n_k\}$ so that

$$|\{x \ : \ |f_{n_k}(x) - f(x)| > 1/k\}| < 2^{-k}.$$

Notice we may do this by the definition of convergence in measure. Without loss of generality, assume that $n_{k+1} > n_k$; if this were not the case, we just drop that and go to the next term. Let

$$E_k = \{x \ : \ |f_{n_k}(x) - f(x)| > 1/k\}.$$

We see that

$$\sum_{k=1}^{\infty} |E_k| < \infty.$$

By Borel-Cantelli (from **Homework 1**), we get that $|\limsup_k E_k| = 0$. But, if $x \in E$, where $f(x) = \lim_{k\to\infty} f_{n_k}(x)$, then

$$|f_{n_k}(x) - f(x)| > \delta$$

for infinitely many k, and in particular we get

$$|f_{n_k}(x) - f(x)| > \delta > 1/k$$

for some eventual k. Thus, we get $x \in \limsup_k E_k$, which has measure zero. Thus, $f_{n_k} \to f$ a.e. **Q.E.D**

Remark 15. This trick is especially important, and is abused often for proving a.e. convergence. We formalize this trick with the next corollary.

Corollary 3.13.1. If for all $\epsilon > 0$ we have

$$\sum_{n=1}^{\infty} |\{f_n - f| > \epsilon\}| < \infty$$

then $f_n \to f$ a.e.

3.6 Lecture 11 (Finishing Results)

We now prove a similar theorem to **Theorem 3.13**.

Theorem 3.14. Let f_n, f be measurable functions which are finite a.e. on E, and $|E| < \infty$. If $f_n \to f$ a.e. then $f_n \xrightarrow{m} f$.

Before proving this, we'd like to establish a basic result.

Lemma 3.6. (i) If $|\bigcup_n E_n| < \infty$ and E_n is a measurable set for all n, then

$$\limsup_{n\to\infty} |E_n| \leqslant \left|\limsup_{n\to\infty} E_n\right|.$$

(ii) If E_n is a measurable set for all n, then

$$\liminf_{n\to\infty} |E_n| \geqslant \left|\liminf_{n\to\infty} E_n\right|.$$

Proof. (i) Let $H_n = \bigcup_{m=n}^\infty E_m$. We see that this forms a decreasing sequence, and moreover by definition we see that $H_n \searrow \limsup E_n$. By **Proposition 2.1 (ii)**, we get

$$\lim_{n\to\infty} |H_n| = \lim_{n\to\infty} \left|\bigcup_{m=n}^\infty E_m\right| = \left|\limsup_{n\to\infty} E_n\right|.$$

Now we have $E_k \subseteq \bigcup_{m=n}^\infty E_m$ for all $k \geqslant m$, so therefore

$$|E_k| \leqslant \left|\bigcup_{m=n}^\infty E_m\right|, \ \forall k \geqslant m.$$

Therefore,

$$\sup_{m\geqslant n} |E_m| \leqslant \left|\bigcup_{m=n}^\infty E_m\right|,$$

and so substituting this in we have

$$\limsup_{n\to\infty} |E_n| \leqslant \left|\limsup_{n\to\infty} E_n\right|.$$

(ii) This is proved analogously. Let $H_n = \bigcap_{m=n}^\infty E_m$. Then we see this forms an increasing sequence, and moreover by definition we see that $H_n \nearrow \liminf E_n$. By **Proposition 2.1 (i)**, we get

$$\lim_{n\to\infty} |H_n| = \lim_{n\to\infty} \left|\bigcap_{m=n}^\infty E_m\right| = \left|\liminf_{n\to\infty} E_n\right|.$$

Now, notice that

$$\bigcap_{m=n}^\infty E_m \subseteq E_k, \ \forall k \geqslant m.$$

Therefore,

$$\left|\bigcap_{m=n}^\infty E_m\right| \leqslant |E_k|, \ \forall k \geqslant m.$$

Thus

$$\left| \bigcap_{m=n}^{\infty} E_m \right| \leqslant \inf_{m \geqslant n} |E_m|.$$

So we get

$$\left| \liminf_{n\to\infty} E_n \right| \leqslant \liminf_{n\to\infty} |E_n|.$$

Q.E.D

Now we may prove the theorem.

Proof. Look at $F_{n,\epsilon} = \{x \;:\; |f_n - f| > \epsilon\}$. Then we need to show that $|F_{n,\epsilon}| \to 0$ as $n \to \infty$ for all $\epsilon > 0$. Using **Lemma 3.6 (i)**, it suffices to then show that

$$\left| \limsup_{n\to\infty} F_{n,\epsilon} \right| = 0.$$

We have

$$\limsup_{n\to\infty} F_{n,\epsilon} = \{x \;:\; |f_n - f| > \epsilon \text{ for infinitely many } n\} \subseteq \{x \;:\; f(x) \neq \lim_{n\to\infty} f_n(x)\}.$$

By assumption, this has measure 0. So therefore we get that the limit of the measure goes to 0. **Q.E.D**

Example 3.7. Here, we give an example that convergence in measure does not imply almost everywhere convergence. Let $n = 2^k + j$ for $j \in \{0, 1, \ldots, 2^{k-1}\}$. Notice that we may write every integer in such a form. Let f_n be a function such that

$$f_n = \chi_{[j/2^k, (j+1)/2^k)}.$$

Then we have

$$f_1 = \chi_{[0,1)},$$
$$f_2 = \chi_{[0,1/2)},\;\; f_3 = \chi_{[1/2,1)},$$
$$f_4 = \chi_{[0,1/4)},\;\; f_5 = \chi_{[1/4,1/2)},\;\; f_6 = \chi_{[1/2,3/4)},\;\; f_7 = \chi_{[3/4,1)}.$$

That is, we have that $f_{2^j}, \ldots, f_{2^{j+1}-1}$ cover the interval $[0,1)$. We first want to see that it converges in measure to 0. That is, we want to show that

$$\lim_{n\to\infty} |\{x \;:\; |f_n - 0| > \epsilon\}| = 0.$$

We may find a j so that

$$|\{x \;:\; |f_{2^j} - 0| > \epsilon\}| < \frac{1}{2^j}$$

for all $\epsilon > 0$ by construction (since $|f_{2^j}| = 1/2^j$). Hence, taking the limit as $n \to \infty$, we get that this goes to 0, and so we get convergence in measure. However, for every $x \in [0,1]$ we have infinitely many n where it is 0 and where it is 1. Therefore, f_n does **not** converge almost everywhere to 0. This is an example of how the points may move around the interval.

We will finish this chapter off by stating a fact which will not be proven in the class due to time.

Theorem 3.15. We have that $f_n \xrightarrow{m}$ if and only if $\lim_{m,n\to\infty} |\{f_n - f_m| > \epsilon\}| = 0$. That is, convergence in measure is equivalent to Cauchy convergence in measure.

Chapter 4

Lebesgue Integration

4.1 Lecture 11 (Lebesgue Integral)

The idea of Lebesgue integration is that we want to define $\int_E f(x)dx = \int_E f$ to be the net volume/area of the region between the graph of f and the x-plane, whatever that may be. We say net because the garph below the plane should be associate with a negative value.

We'll start by only dealing with non-negative functions.

Definition. We define the **graph of f over E** to be

$$\Gamma(f, E) = \{(x, f(x)) \ : \ x \in E\}.$$

Definition. We define the **region underneath a function over E** to be

$$R(f, E) = \{(x, y) \ : \ x \in E, \ 0 \leqslant y \leqslant f(x) \text{ if } f < \infty, \text{ or } 0 \leqslant y < \infty \text{ if } f(x) = \infty\}.$$

Definition. We define the **Lebesgue integral of f over E**, denoted by $\int_E f$, to be $|R(f, E)|$ if this region is measurable.

The next theorem connects our study of measurable functions (**Chapter 3**) with integrability.

Theorem 4.1. If $f \geqslant 0$ on a measurable set E, then $\int_E f$ is defined if and only if f is measurable.

Before proving this, we first look at an important example.

Example 4.1. Let

$$f = \sum_{i=1}^{n} a_i \chi_{E_i} \geqslant 0$$

be a simple measurable function, and $E = \bigcup_{i=1}^{n} E_i$. We examine then

$$R(f, E) = \bigcup_{i=1}^{n} (E_i \times [0, a_i]) \, .$$

Since the E_i are disjoint by assumption, we may write this as

$$|R(f,E)| = \sum_{i=1}^{n} |E_i||[0,a_i]| = \sum_{i=1}^{n} |E_i|a_i = \int_E f.$$

This gives us that simple measurable functions are integrable.

We will also need a somewhat easy lemma and a not so easy lemma before proving the theorem. Following is the somewhat easy lemma.

Lemma 4.1. If f is a measurable function over E, $f \geqslant 0$, then we have $|\Gamma(f,E)| = 0$.

Proof. Fix $\epsilon > 0$ and let $E_k = \{x \ : \ k\epsilon \leqslant f(x) < (k+1)\epsilon\}$. In other words, we cut everything up into strips and collect the points which map into those strips. We look now at

$$\Gamma(f,E) \subseteq \bigcup_k \left(E_k \times [k\epsilon, (k+1)\epsilon)\right).$$

Therefore, we have

$$|\Gamma(f,E)|_e \leqslant \sum_k |E_k||\epsilon| = |\epsilon| \sum_k |E_k|.$$

Since the E_k are disjoint, we get

$$\sum_k |E_k| = |E|.$$

Thus, we have

$$|\Gamma(f,E)|_e \leqslant |E||\epsilon|.$$

If $|E| < \infty$, we have then that $|\Gamma(f,E)|_e = 0$ by just taking the infimum over all ϵ. Otherwise, we must examine $|\Gamma(f, E \cap \{|x| \leqslant n\}|_e$. By our prior work, we see that this is 0 for all n, and since thsi goes to $\Gamma(f,E)$ as we let $n \to \infty$ and this is an increasing sequence we get that $|\Gamma(f,E)| = 0$. **Q.E.D**

Following this is the not so easy lemma.

Lemma 4.2. If $A \subseteq \mathbb{R}^n$ is such that $A \times [a,b]$ is measurable for some $a < b$, then A is measurable.

Proof. We prove this in a series of steps, following the proof found here.

(Step 1) For any $A \subseteq \mathbb{R}^n$, $|A \times [0,1]|_e \leqslant |A|_e$.

Proof. For any $\epsilon > 0$, take an cover of intervals $A \subseteq \bigcup_k I_k$ such that $\sum_k |I_k| < |A|_e + \epsilon$. Then we have

$$|A \times [0,1]|_e \leqslant \sum_k |I_k \times [0,1]| = \sum_k |I_k| < |A|_e + \epsilon.$$

Since this works for any ϵ we get the results. **Q.E.D**

(Step 2) If $A \times [0,1]$ is measurable, then for any $\epsilon > 0$ there exists an open set G containing A such that $(G \backslash A) \times [0,1]$ is measurable and $|(G \backslash A) \times [0,1]| < \epsilon$.

Proof. If $A \times [0,1]$ is measurable, then we have that we may pick an open $H \subseteq \mathbb{R}^{n+1}$ so that $A \times [0,1] \subseteq H$ and $|H \backslash (A \times [0,1])| < \epsilon$. Fixing an $x \in A$, the function $y \mapsto \text{dist}((x,y), H^c)$ is a continuous, positive function on $[0,1]$, and so there is a positive minimum which we denote by $\delta_x > 0$. Therefore,

$$(x - \delta_x, x + \delta_x) \times [0,1] \subseteq H$$

for all $x \in A$. Letting

$$G = \bigcup_{x \in A} (x - \delta_x, x + \delta_x),$$

we have that G is an open set which contains A such that $G \times [0,1] \subseteq H$. Since G and A are measurable, $(G \backslash A) \times [0,1] = (G \times [0,1]) \backslash (A \times [0,1])$ is also measurable. Furthermore,

$$|(G \backslash A) \times [0,1]| \leqslant |H \backslash (A \times [0,1])| < \epsilon.$$

Q.E.D

(Step 3) If $A \times [0,1]$ is measurable, then $|A \times [0,1]| = |A|_e$.

Proof. Take G open as in **Step 2**. Then we get that, for all $\epsilon > 0$,

$$|A \times [0,1]| = |G \times [0,1]| - |(G \backslash A) \times [0,1]| > |G \times [0,1]| - \epsilon,$$

per **Step 2**. Using the fact that $|G \times [0,1]| = |G|$, we then get that

$$|A \times [0,1]| > |G| - \epsilon.$$

Since $A \subseteq G$, we get that $|A|_e \leqslant |G|$, so that

$$|A \times [0,1]| > |A|_e - \epsilon.$$

This works for all ϵ, so taking the supremum gives

$$|A|_e \leqslant |A \times [0,1]|.$$

By **Step 1**, we get equality. **Q.E.D**

(Step 4) If $A \times [0,1]$ is measurable, then A is measurable.

Proof. Taking G again as in **Step 2**, we have for all $\epsilon > 0$

$$|G \backslash A|_e = |(G \backslash A) \times [0,1]| < \epsilon.$$

By definition we win. **Q.E.D**

This then concludes the proof. **Q.E.D**

We now prove the theorem.

Proof. ($\Longleftarrow$) Assume f is measurable, $f \geqslant 0$. By **Theorem 3.9**, we get that there is a sequence of $f_n \nearrow f$ almost everywhere. For simple functions, we have that $R(f_n, E)$ is measurable by **Example 4.1**. This almost approaches $R(f, E)$; however, it could be that this misses the actual graph of f. Thus, throwing this in, we get

$$R(f_n, E) \cup \Gamma(f, E) \nearrow R(f, E).$$

Since $\Gamma(f, E)$ has measure 0 by **Lemma 4.1** and the $R(f_n, E)$ are all measurable, we get that $R(f, E)$ is measurable. ($\Longrightarrow$) It suffices to prove that $\{f > 0\}$ is measurable, since measurability is preserved under shifts. To see that this is measurable, let

$$R_1(f, E) = \{(x, y) \ : \ x \in E, \ 0 < y \leqslant f(x)\}.$$

Since R_1 is measurable (by assumption), we get that any vertical stretch of R_1 is also measurable, since a stretch is just a Lipschitz transformation. More precisely, we have for $n \geqslant 1$ that

$$R_n = \{(x, y) \ : \ x \in E, \ 0 < y \leqslant nf(x)\}$$

is measurable. We get that $R_n \nearrow \{f > 0\} \times (0, \infty)$, and so this set is measurable. Intersecting this with $\mathbb{R}^n \times [1, 2]$, we get that $\{f > 0\} \times [1, 2]$ is measurable. By **Lemma 4.2**, this gives us that $\{f > 0\}$ is measurable. **Q.E.D**

4.2 Lecture 12 (Non-Negative Lebesgue Integral Properties)

Theorem 4.2. (Properties of Lebesgue Integral) Assume throughout that all functions are measurable and non-negative.

(i) If $0 \leqslant g \leqslant f$ then $\int_E g \leqslant \int_E f$.

(ii) If $\int_E f < \infty$, then $f < \infty$ a.e. on E.

(iii) If $E_1 \subseteq E_2$, then $\int_{E_1} f \leqslant \int_{E_2} f$.

(iv) If $E = \bigcup_{j=1}^{\infty} E_j$, E_j disjoint, then

$$\int_E f = \sum_{j=1}^{\infty} \int_{E_j} f.$$

(v) If $|E| = 0$, then $\int_E f = 0$.

(vi) If $g \leqslant f$ a.e. on E, then $\int_E g \leqslant \int_E f$.

(vii) If $f = g$ a.e. on E, then $\int_E g = \int_E f$.

(viii) Assuming $|E| > 0$, we have $\int_E f = 0$ if and only if $f = 0$ a.e. on E.

(ix) Integration is linear; that is, for $\alpha, \beta \in \mathbb{R}$, we have

$$\int_E (\alpha f + \beta g) = \alpha \int_E f + \beta \int_E g.$$

Before proving the theorem, we will state some important theorems that we will use along the way.

Theorem 4.3. (Monotone Convergence Theorem) If $\{f_n\}$ is a sequence of non-negative functions such that $f_n \nearrow f$ a.e. on E, then $\lim_{n\to\infty} \int_E f_n = \int_E f$.

Proof. Examine $R(f_n, E)$. Since f_n is increasing, then $R(f_n, E)$ is increasing to $R(f, E)$, except we may miss $\Gamma(f, E)$. Thus, we have $R(f_n, E) \cup \Gamma(f, E) \nearrow R(f, E)$. Now, using the definition of integral and the fact that $|\Gamma(f, E)| = 0$ from **Lemma 4.1**, we have

$$\int_E f = |R(f, e)| = \lim_{n\to\infty} |R(f_n, E) \cup \Gamma(f, E)| = \lim_{n\to\infty} |R(f_n, E)| = \lim_{n\to\infty} \int_E f_n.$$

Q.E.D

Theorem 4.4. (Chebychev's Theorem) We have for measurable non-negative f that

$$|\{x \in E \ : \ f > a\}| \leqslant \frac{1}{a} \int_E f.$$

Proof. We may bound $\int_E f$ below by

$$\int_{E \cap \{f > a\}} f \leqslant \int_E f.$$

We may bound this below again by the constant function a; that is, we have the chain

$$\int_{E \cap \{f > a\}} a \leqslant \int_{E \cap \{f > a\}} f \leqslant \int_E f.$$

Solving the integral on the far left, we have

$$\int_{E \cap \{f > a\}} a = a \cdot |\{x \in E \ : \ f > a\}|.$$

So rewriting everything we have

$$a \cdot |\{x \in E \ : \ f > a\}| \leqslant \int_E f.$$

Dividing throughout by a gives the desired result. **Q.E.D**

We now prove **Theorem 4.2**.

Proof. (i) This follows clearly, since $R(g,E) \subseteq R(f,E)$ and so $|R(g,E)| \leqslant |R(f,E)|$.

(ii) We may bound the integral below by

$$|\{f = \infty\} \times [0,\infty)| \leqslant \int_E f < \infty.$$

Now notice we may write the left hand side as

$$|\{f = \infty\} \times [0,\infty)| = |\{f = \infty\}| \cdot \infty.$$

Since we require this to be finite, the only way we can get that is if $|\{f = \infty\}| = 0$. So, we get that $f < \infty$ a.e. on E.

(iii) Again, we have $R(f,E_1) \subseteq R(f,E_2)$ so that $|R(f,E_1)| \leqslant |R(f,E_2)|$.

(iv) We may write

$$R(f,E) = \bigcup_{j=1}^{\infty} R(f,E_j).$$

Since the E_j are disjoint, the $R(f,E_j)$ are also disjoint. Thus, taking it's measure, we get

$$\left|\bigsqcup_{j=1}^{\infty} R(f,E_j)\right| = \sum_{j=1}^{\infty} |R(f,E_j)| = \sum_{j=1}^{\infty} \int_{E_j} f.$$

(v) We have $R(f,E) \subseteq E \times [0,\infty)$. Thus, $|R(f,E)| \leqslant |E| \cdot \infty = 0$. So $|R(f,E)| = 0$.

(vi) We may write

$$\int_E g = \int_{E\cap\{f<g\}} g + \int_{E\cap\{f\geqslant g\}} g.$$

Now, since $f \geqslant g$ almost everywhere, we have that $E \cap \{f < g\}$ has measure 0. Therefore, we get

$$\int_E g = \int_{E\cap\{f\geqslant g\}} g \leqslant \int_E f,$$

as desired.

(vii) We use the prior property, with the inequality going both ways.

(viii) ($\Longleftarrow$) This follows by the prior property.
($\Longrightarrow$) We use **Theorem 4.4** here. Since $0 = \int_E f$, we have

$$|\{x \in E \ : \ f > a\}| \leqslant \frac{1}{a}\int_E f = 0.$$

Now, for a decreasing to 0, we get $\{x \in E \ : \ f > a\} \nearrow \{x \in E \ : \ f > 0\}$. Therefore, we have that $|\{x \in E \ \ f > 0\}| = 0$.

(ix) We prove this in two steps.

(Step 1) For c constant, we have $\int_E cf = c\int_E f$.

Proof. The transformation $L(x,y) = (x, cy)$ is a linear transformation, which can be represented by the matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix},$$

and so $\det(L) = c$. So, applying this to the set $R(f,E)$, we have $R(cf,E)$. By **Theorem 3.2** we get

$$|R(cf,E)| = c|R(f,E)|.$$

Q.E.D

(Step 2) For f, g non-negative measurable functions, we have $\int_E(f+g) = \int_E f + \int_E g$.

Proof. We first prove this for simple functions. Let f and g be simple measurable functions. Then by definition, we may represent them by

$$f = \sum_{i=1}^{n} a_i \chi_{A_i}, \quad g = \sum_{j=1}^{m} b_j \chi_{B_j},$$

where $E = \sqcup_{i=1}^{n} A_i = \sqcup_{j=1}^{m} B_j$. Therefore, we have

$$f + g = \sum_{i=1}^{n}\sum_{j=1}^{m} (a_i + b_j)\chi_{A_i \cap B_j}.$$

Integrating this over E, we get

$$\int_E (f+g) = \sum_{i=1}^{n}\sum_{j=1}^{m} (a_i + b_j)|A_i \cap B_j|.$$

Now, expanding and using the finite Fubini theorem grants us

$$\sum_{i=1}^{n}\sum_{j=1}^{m} (a_i + b_j)|A_i \cap B_j| = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i|A_i \cap B_j| + \sum_{j=1}^{m}\sum_{i=1}^{n} b_j|A_i \cap B_j|.$$

$$= \sum_{i=1}^{n} a_i \sum_{j=1}^{m} |A_i \cap B_j| + \sum_{j=1}^{m} b_j \sum_{i=1}^{n} |A_i \cap B_j|.$$

Since the A_i and B_j are disjoint and cover the space, we see that summing $|A_i \cap B_j|$ over j gives us just $|A_i|$, and summing it over i gives us just $|B_j|$. Hence, we get

$$\sum_{i=1}^{n} a_i \sum_{j=1}^{m} |A_i \cap B_j| + \sum_{j=1}^{m} b_j \sum_{i=1}^{n} |A_i \cap B_j| = \sum_{i=1}^{n} a_i |A_i| + \sum_{j=1}^{n} b_j |B_j|$$

$$= \int_E f + \int_E g.$$

So for simple measurable non-negative functions, we get the desired result.

Now, take f and g to just be non-negative measurable functions. Then by **Theorem 3.9**, we get that there are sequences of simple measurable non-negative functions $\{f_n\} \nearrow f$ and $\{g_n\} \nearrow g$. We use **Theorem 4.3** to get

$$\int_E (f+g) = \lim_{n\to\infty} \int_E (f_n + g_n).$$

We showed that linearity holds for simple measurable non-negative functions, so we have

$$\lim_{n\to\infty} \int_E (f_n + g_n) = \lim_{n\to\infty} \int_E f_n + \lim_{n\to\infty} \int_E g_n.$$

Using **Theorem 4.3** again, we get

$$\lim_{n\to\infty} \int_E f_n + \lim_{n\to\infty} \int_E g_n = \int_E f + \int_E g.$$

Thus, we have the desired result. **Q.E.D**

(Step 1) and (Step 2) in conjunction give us the desired result. **Q.E.D**

These properties give us the following corollary.

Corollary 4.4.1. If $f_n \geqslant 0$ on E, then

$$\int_E \sum_{n=0}^{N} f_n = \sum_{n=0}^{N} \int_E f_n.$$

Proof. This follows by the linearity property of the Lebesgue integral. **Q.E.D**

However, we can actually use these in conjunction with the prior theorems to get that this holds for even countable sums.

Corollary 4.4.2. If $f_n \geqslant 0$ on E, then

$$\int_E \sum_{n=0}^{\infty} f_n = \sum_{n=0}^{\infty} \int_E f_n.$$

Proof. Let $F_N = \sum_{n=0}^{N} f_n$. Then we have $F_N \nearrow F = \sum_{n=0}^{\infty} f_n$. By **Theorem 4.3**, we have

$$\lim_{N\to\infty} \int_E F_N = \int_E F.$$

Notice that we may rewrite the left hand side to be

$$\lim_{N\to\infty} \int_E \sum_{n=0}^{N} f_n.$$

The sum on the inside of the integral is a finite sum, and so by **Corollary 4.4.1** we have

$$\lim_{N\to\infty} \sum_{n=0}^{N} \int_E f_n = \sum_{n=0}^{\infty} \int_E f_n.$$

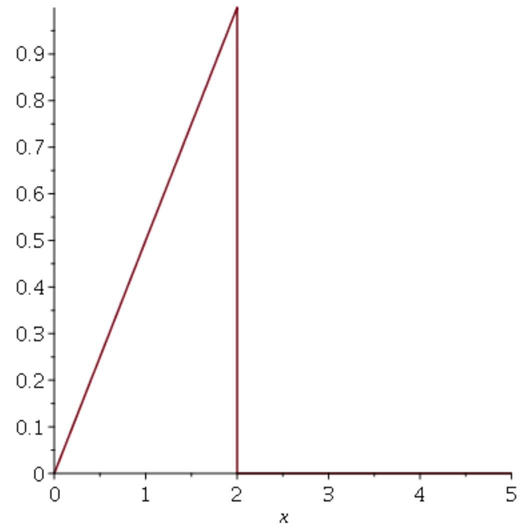
Thus, we have the desired result. **Q.E.D**

We now move on to discuss a little about integration limit theorems. The Monotone Convergence Theorem, **Theorem 4.3**, tells us that we can sometimes pull the limit inside of an integral and things work out fine. One may ask whether this always holds, and the answer is a resounding "Not always." Below are some examples where this does not hold.

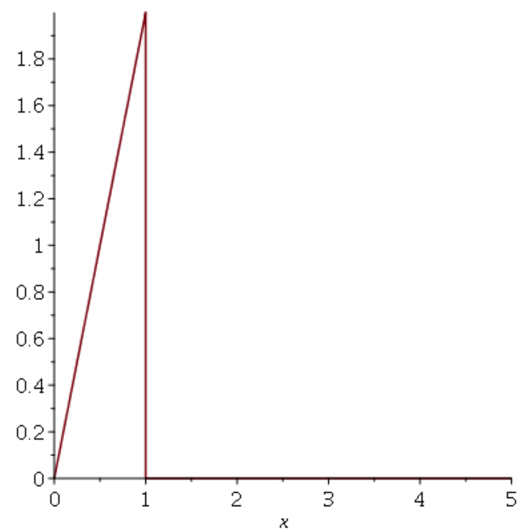
Example 4.2. (a) Let

$$f_n(x) = \begin{cases} \frac{n^2}{2} x \text{ for } 0 \leqslant x \leqslant \frac{2}{n}, \\ 0 \text{ otherwise.} \end{cases}$$

That is, we have a triangle with base $[0, 2/n]$ and height n. We have f_1 plotted below:



and f_2 plotted below:



So we see as n gets larger we get a thinner, taller triangle. We see for all x that for n large enough we will have $f_n(x) = 0$, so we have that $f_n \to 0$ as $n \to \infty$. However, we have that

$$\int_0^{2/n} \frac{n^2}{2} x dx = 1$$

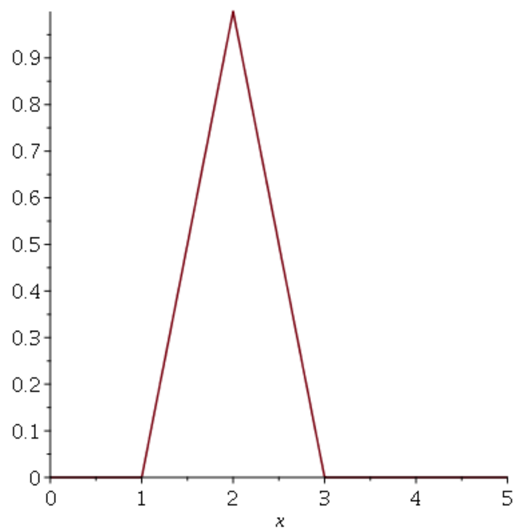
for all n. Therefore, we have

$$\lim_{n\to\infty} \int_E f_n = 1 \neq \int_E \lim_{n\to\infty} f_n = 0.$$

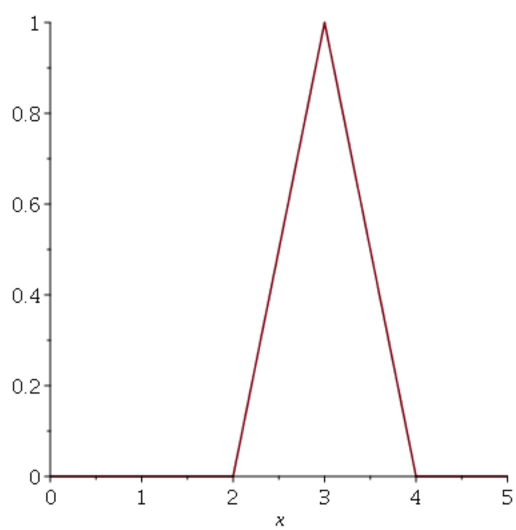
(b) Let

$$f_n(x) = \begin{cases} x - n \text{ for } n \leqslant x \leqslant n+1 \\ -x + n + 2 \text{ for } n+1 \leqslant x \leqslant n+2 \\ 0 \text{ otherwise.} \end{cases}$$

The plot of f_1 is given below:



The plot of f_2 is given below:



So we see that it is a triangle of area 1 moving along the x axis. Again, for any fixed x, we see that $f_n(x) = 0$ for n sufficiently large. So we get

$$\lim_{n\to\infty} \int_E f_n = 1 \neq \int_E \lim_{n\to\infty} f_n = 0$$

again.

We state some theorems which will be proven in the next lecture. These theorems give us some idea of when we are allowed to pull in the limit.

Theorem 4.5. (Fatou's Lemma) If $f_n \geqslant 0$,

$$\int_E \left(\liminf_{n\to\infty} f_n\right) \leqslant \liminf_{n\to\infty} \int_E f_n.$$

Theorem 4.6. (Dominated Convergence Theorem) If $f_n \geqslant 0$, $f_n \to f$ a.e. and there exists a $\phi \geqslant 0$ such that

(i) $f_n \leqslant \phi$ a.e.,

(ii) $\int_E \phi < \infty$,

then

$$\lim_{n\to\infty} \int_E f_n = \int_E f.$$

4.3 Lecture 13 (Non-Negative Fatou, General Lebesgue Integral)

We will first prove Fatou's Lemma.

Proof. Let $g_n = \inf_{m \geqslant n} f_m$. Then we see that $g_n \nearrow \liminf f_n$ by definition. We have that **Theorem 4.3** tells us that

$$\int_E \liminf_{n\to\infty} f_n = \lim_{n\to\infty} \int_E g_n.$$

Notice that $f_n \geqslant g_n$ for all n. Then **Theorem 4.2 (i)** tells us that

$$\int_E f_n \geqslant \int_E g_n,$$

and in particular

$$\liminf_{n\to\infty} \int_E f_n \geqslant \lim_{n\to\infty} \int_E g_n.$$

Chaining these inequalities together gives us the desired result. **Q.E.D**

It turns out that the Dominated Convergence Theorem is just a consequence of Fatou's Lemma. We prove this now.

Proof. Fatou gives us half of what we want. That is, since we have almost everywhere convergence, we get that

$$\liminf_{n\to\infty} f_n = \lim_{n\to\infty} f_n = f \text{ a.e.}$$

and so by Fatou's we have

$$\int_E f \leqslant \liminf_{n\to\infty} \int_E f.$$

In order to get the desired result, we must then show

$$\limsup_{n\to\infty} \int_E f_n \leqslant \int_E f.$$

Let $g_n = \phi - f_n$. Then we have that $g_n \geqslant 0$, and by Fatou again we have

$$\int_E \liminf_{n\to\infty} g_n \leqslant \liminf_{n\to\infty} \int_E g_n.$$

Writing out, this is

$$\int_E (\phi - f) \leqslant \liminf_{n\to\infty} \int_E (\phi - f_n).$$

Linearity of integrals almost gives us what we want. However, we only have this for addition. We can cheat this by setting

$$\phi = (\phi - f) + f.$$

Then we have

$$\int_E \phi = \int_E (\phi - f) + \int_E f$$

by linearity. Since $\int_E \phi$ is finite, we can subtract $\int_E f$ (which must also be finite) to get

$$\int_E \phi - \int_E f = \int_E (\phi - f).$$

So, rewriting the above inequality using this result, we have

$$\int_E \phi - \int_E f \leqslant \liminf_{n\to\infty} \left(\int_E \phi - \int_E f \right).$$

Now, distributing the limit, we have

$$\int_E \phi - \int_E f \leqslant \int_E \phi + \left(\liminf_{n\to\infty} - \int_E f_n \right) = \int_E \phi - \limsup_{n\to\infty} \int_E f_n.$$

Using the finiteness of everything, we are allowed to rearrange this to get

$$\limsup_{n\to\infty} \int_E f_n \leqslant \int_E f.$$

Q.E.D

The next question we want to explore is how to integrate general functions. The idea is to break it down into it's non-negative components and try to use prior results on that. This leads us to our first definition.

Definition. For general functions, we write

$$f = f^+ - f^-.$$

We thus set

$$\int_E f = \int_E f^+ - \int_E f^-,$$

as long as either $\int_E f^+ < \infty$ or $\int_E f^- < \infty$.

We now can talk about the Lebesgue integral for general functions.

Definition. We say that a function f is **Lebesgue integrable on E**, denoted by $f \in L(E)$, if $\int_E f$ exists and is finite.

Remark 16. We may have that the integral exists, even though the function itself is not integrable.

Lemma 4.3. We have that $f \in L(E)$ if and only if $|f| \in L(E)$.

Proof. Note that we define $|f| = f^+ + f^-$. Assume that f integrable. Then we have

$$\int_E f = \int_E f^+ - \int_E f^- < \infty.$$

This implies that we must have that $\int_E f^+$ and $\int_E f^-$ are both finite. Likewise, if

$$\int_E |f| = \int_E f^+ + \int_E f^- < \infty,$$

then both $\int_E f^+$ and $\int_E f^-$ exist and are finite. **Q.E.D**

Lemma 4.4. We have

$$\left|\int_E f\right| \leqslant \int_E |f|.$$

Proof. We see that

$$\left|\int_E f\right| = \left|\int_E f^+ - \int_E f^-\right|.$$

By the triangle inequality, we get

$$\left|\int_E f^+ - \int_E f^-\right| \leqslant \left|\int_E f^+\right| + \left|\int_E f^-\right| = \int_E |f|.$$

Q.E.D

We now get a theorem of properties of the Lebesgue integral, which is similar to that of **Theorem 4.2**. We will omit the proof of most of these as a result.

Theorem 4.7. (i) We have that $f \in L(E)$ implies $|f| < \infty$ almost everywhere on E.

(ii) If $f \leqslant g$ almost everywhere and $f, g \in L(E)$, then $\int_E f \leqslant \int_E g$.

(iii) If $E_1 \subseteq E_2$ and $f \in L(E_2)$, then $f \in L(E_1)$ as well (**Notice no inequality**).

(iv) If $f \in L(E)$ and $E = \bigsqcup_k E_k$, then

$$\int_E f = \sum_k \int_{E_k} f,$$

and moreover $f \in L(E_k)$ for each k.

(v) $|E| = 0$ implies $\int_E f = 0$.

(vi) $f = 0$ almost everywhere in E implies that $\int_E f = 0$.

(vii) If $f \in L(E)$, then $cf \in L(E)$, and moreover

$$\int_E cf = c \int_E f.$$

(viii) If $f, g \in L(E)$, then $f + G \in L(E)$ and moreover

$$\int_E (f + g) = \int_E f + \int_E g.$$

Proof. We include the proof for the last property. First, notice that $f+g \in L(E)$, since $|f + g| \leqslant |f| + |g|$ by the triangle inequality. By **Lemma 4.3**, the first property, and **Theorem 4.2**, we get

$$\int_E |f| + \int_E |g| < \infty \implies \int_E |f + g| < \infty,$$

and since $|f+g| \in L(E)$ we have $f+g \in L(E)$. Now, we show the second part. If $f, g \geqslant 0$, we win. Let's show it holds if $f \geqslant 0$, $g < 0$, and $f + g \geqslant 0$ everywhere. Write $f = (f + g) + (-g)$. Then this implies that we have

$$\int_E f = \int_E (f+g) + \int_E (-g) = \int_E (f+g) - \int_E g \implies \int_E f + \int_E g = \int_E (f+g),$$

since things are finite. Abusing multiplication by -1, we see that we have that as long as f, g and $f + g$ have constant sign over E, then we win. But we can just divide up E into a disjoint union of E_k, $k \in \{1, \ldots, 6\}$, where in each E_k we have that they have constant sign. Therefore, we have

$$\int_E (f+g) = \sum_{k=1}^{6} \int_{E_k} (f+g) = \sum_{k=1}^{6} \left(\int_{E_k} f + \int_{E_k} g \right) = \sum_{k=1}^{6} \int_{E_k} f + \sum_{k=1}^{6} \int_{E_k} g$$
$$= \int_E f + \int_E g.$$

Q.E.D

We'd like to now translate the limit theorems to statement about general functions.

Theorem 4.8. (MCT revised)

(i) If $f_n \nearrow f$ and $\phi \leqslant f_n$ with $\phi \in L(E)$, then $\int_E f_n \to \int_E f$.

(ii) If $f_n \searrow f$ and $f_n \leqslant \phi$ with $\phi \in L(E)$, then $\int_E f_n \to \int_E f$.

Proof. We prove (i). The proof of (ii) is just (i) but multiplied by -1. Examing $g_n = f_n - \phi$. By the old MCT, we have that

$$\int_E g_n \to \int_E g.$$

Rewriting this, we have

$$\int_E f_n - \phi \to \int_E f - \phi,$$

and so by linearity we get

$$\int_E f_n \to \int_E f.$$

Q.E.D

Theorem 4.9. (Fatou's Lemma revised)

(i) If there exists a $\phi \in L(E)$ such that $f_n \geqslant \phi$ for all n, then

$$\int_E \liminf_{n\to\infty} f_n \leqslant \liminf_{n\to\infty} \int_E f_n.$$

(ii) If there exists a $\phi \in L(E)$ such that $f_n \leqslant \phi$ for all n, then

$$\int_E \limsup_{n\to\infty} f_n \geqslant \limsup_{n\to\infty} \int_E f_n.$$

Proof. (i) Apply the old Fatou's lemma (**Theorem 4.5**) to $f_n - \phi$ and use linearity.

(ii) Use the proof of **Theorem 4.6**, noting that

$$\liminf_{n\to\infty} -f_n = -\limsup_{n\to\infty} f_n.$$

Q.E.D

Theorem 4.10. (DCT revised) If $f_n \to f$ a.e. on E and $\sup_n |f_n| \leqslant \phi \in L(E)$, then

$$\int_E f_n \to \int_E f.$$

Proof. This is a consequence of **Theorem 4.9**, using that $-\phi \leqslant f_n \leqslant \phi$ and then using parts (i) and (ii). **Q.E.D**

4.4 Lecture 14 (BCT, UCT, Riemann Integral)

We want to talk about two more limit theorems which are extremely useful. For these, we will note that we need another assumption; that is, we need $|E| < \infty$.

Theorem 4.11. (Bounded Convergence Theorem) If $|E| < \infty$, $f_n \to f$ almost everywhere on E, and $|f_n(x)| \leqslant M$ for all $x \in E$ and for all $n \geqslant 1$, then

$$\int_E f_n \to \int_E f.$$

Proof. This is actually a corollary of **Theorem 4.6**, because $\phi = M$ is integrable on a set of finite measure. **Q.E.D**

Theorem 4.12. (Uniform Convergence Theorem) If $|E| < \infty$, $f_n \to f$ uniformly on E, and $f_n \in L(E)$, then

$$\int_E f_n \to \int_E f,$$

and furthermore $f \in L(E)$.

Proof. Recall that $f_n \to f$ uniformly means that for all $\epsilon > 0$ there exists an n_0 such that for all $n \geqslant n_0$,

$$|f_n(x) - f(x)| < \epsilon, \ \forall \epsilon > 0.$$

Notice that we may write

$$f(x) = f(x) - f_n(x) + f_n(x).$$

By the triangle inequality, we have

$$|f(x)| \leqslant |f(x) - f_n(x)| + |f_n(x)|.$$

Taking n sufficiently large, we have that this is bounded above by

$$|f(x)| \leqslant \epsilon + |f_n(x)|.$$

Integrating both sides gives us

$$\int_E |f| \leqslant \int_E \epsilon + \int_E |f_n| = |E|\epsilon + \int_E |f_n|$$

Since the right hand side is finite (assuming $f_n \in L(E)$ and $|E| < \infty$), we get that the left hand side is finite. So $|f| \in L(E)$, which by **Lemma 4.3** implies $f \in L(E)$. This gives us the second part. For the first part, we want to prove that

$$\int_E f_n \to \int_E f,$$

or, in other words,

$$\left| \int_E f_n - \int_E f \right| \to 0.$$

By the linearity of integration, the inside may be rewritten as

$$\left| \int_E (f_n - f) \right|.$$

By the proof of **Theorem 4.7**, we get

$$\left| \int_E (f_n - f) \right| \leqslant \int_E |f_n - f|.$$

Now, since we have uniform convergence, for fixed ϵ we know that for n sufficiently large we have $|f_n - f| < \epsilon$ uniformly over E. So we get

$$\int_E |f_n - f| \leqslant \int_E \epsilon = |E|\epsilon.$$

Since this applies for all $\epsilon > 0$, we take $\epsilon \to 0$ to get

$$\int_E |f_n - f| \leqslant 0,$$

as desired. **Q.E.D**

We now want to relate Lebesgue and Riemann integrals. For notations sake, we will denote the Riemann integral by

$$(R) \int_a^b f = \lim_{\text{mesh}(\pi_n) \to 0} \sum_i f(x_i^*)(x_i - x_{i+1}),$$

where π_n denotes some partition

$$a = x_0 < x_1 < \ldots < x_n = b,$$

$x_i^* \in [x_{i-1}, x_i]$, and $\text{mesh}(\pi_n)$ denotes the max distance between two consecutive points in the partition.

Theorem 4.13. Let f be bounded on $[a, b]$. If f is Riemann integrable on $[a, b]$, then $f \in L([a, b])$ and

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

Remark 17. The assumption that f bounded is **not** necessary, as we will see towards the end.

Proof. For $n \geqslant 1$, let π_n be the partition $\{x_i\}_{i=0}^{2^n}$. That is, the partition of Dyadic intervals. More specifically, we have that

$$x_k = a + \frac{k}{2^n}(b-a).$$

Now, let

$$l_n(x) := \sum_{k=1}^{2^n} \left(\inf_{[x_{k-1},x_k]} f(x) \right) \chi_{[x_{k-1},x_k]}(x),$$

$$u_n(x) := \sum_{k=1}^{2^n} \left(\sup_{[x_{k-1},x_k]} f(x) \right) \chi_{[x_{k-1},x_k]}(x),$$

$$L_n = \int_a^b l_n(x),$$

$$U_n = \int_a^b u_n(x),$$

where the L_n and U_n are Lebesgue integrals. The assumption was that f was Riemann integrable, which gives us that $L - n$ and U_n converge to the same limit as $n \to \infty$;

$$L_n \to (R) \int_a^b f,$$

$$U_n \to (R) \int_a^b f.$$

So all we need to prove is that these converge as Lebesgue integrals as well.

Let's next note that l_n forms an increasing sequence pointwise, while u_n forms a decreasing sequence pointwise, since we are taking subdivisions. We see that they are *probably* increasing to f, but we can't say that for certain quite yet. For now, denote these limits as l and u respectively. The BCT (**Theorem 4.11**) tells us that

$$\int_a^b l_n \to \int_a^b l,$$

$$\int_a^b u_n \to \int_a^b u.$$

So we have

$$(R) \int_a^b f = \int_a^b l = \int_a^b u.$$

Note that, by construction, $u - l \geqslant 0$. Then we have

$$\int_a^b u - l = 0,$$

which implies that $u - l = 0$ almost everywhere. This is the same as saying $y = k$ almost everywhere on $[a, b]$. So since $l \leqslant f \leqslant u$, we get $l = f = u$ almost everywhere on $[a, b]$. So

$$\int_a^b l = \int_a^b f = \int_a^b u$$

by prior results. **Q.E.D**

This proof also gives us a nice characterization of Riemann integrals.

Theorem 4.14. Assume f is bounded on $[a, b]$. Then f is Riemann integrable on $[a, b]$ if and only if f is continuous a.e. on $[a, b]$.

Proof. ($\Longrightarrow$) Assume it is Riemann integrable. Let l_n, u_n, l, u be as above. Then we know $l = f = u$ almost everywhere. Say $Z_0 = \{x \in [a, b] \ : \ l(x) \neq u(x)\}$. Then since it's almost everywhere, we have $|Z_0| = 0$. Let $Z_1 = \{x \in [a, b] \ : \ x = a + \frac{b}{2^n}(b-a), \ n \geqslant 1, \ k \leqslant 2^n\}$. Then $|Z_1| = 0$ since it is a countable collection of points. Let $Z = Z_0 \cup Z_1$. Then by construction $|Z| = 0$. We then claim that if $x \notin Z$, w ehave f is continuous at x. Fix an $\epsilon > 0$, and pick n large enough so that $l_n(x) > f(x) - \epsilon$, $u_n(x) < f(x) + \epsilon$. We may do this since $x \notin Z_0$. For this n fixed, we have $x \in (x_{k-1}^{(n)}, x_k^{(n)})$, where this is an open interval since $x \notin Z_1$. For y in this interval, we know that $|f(x) - f(y)| < \epsilon$, since l_n is an infimum and u_n a supremum. Thus, we have found an appropriate interval, and so it is continuous a.e.
($\Longleftarrow$) Assume f is continuous a.e Let π_n be some sequence of partitions with the mesh going to 0. Let l_n, u_n be the corresponding step function for the upper and lower Riemann sums. If f is continuous at x, then $l_n(x) \to f(x)$ and $u_n \to f(x)$ a.e. Therefore, by the BCT, we have

$$\int_a^b l_n \to \int_a^b f,$$

$$\int_a^b u_n \to \int_a^b f.$$

Since these are the same, we get that f is Riemann integrable. **Q.E.D**

We end by talking about improper Riemann integrals.

Remark 18. Note that there is no need for "improper" Lebesgue integrals, by definition.

Theorem 4.15. Assume $f \geqslant 0$ on $[a, b]$ and f is Riemann integrable (and therefore bounded) on $[a + \epsilon, b]$ for all $\epsilon > 0$. If

$$\lim_{\epsilon \to 0} (R) \int_{a+\epsilon}^b f = I,$$

then $f \in L([a, b])$ and

$$\int_a^b f = I.$$

Proof. We sketch the proof her. Let $f_\epsilon(x) = f(x)\chi_{[a+\epsilon,b]}$. Note that $f_\epsilon(x) \nearrow f(x)\chi_{(a,b]}(x)$ as $\epsilon \to 0$. Then

$$(R)\int_{a+\epsilon}^{b} f = \int f_\epsilon \to \int_a^b f$$

by the MCT on $[a,b]$. **Q.E.D**

Remark 19. Here, note that the $f \geqslant 0$ assumption **is** necessary.

4.5 Lecture 15 (Fubini's Theorem)

The question we'd like to answer today is, assuming $f(x,y)$ is a measurable function on $\mathbb{R}^{n+m}$, where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, when is

$$\int f(x,y)(dxdy) = \int\left(\int f(x,y)dx\right)dy = \int\left(\int f(x,y)dy\right)dx.$$

That is, when can we take an iterated integral? To answer this question, we must consider two other sub-questions: for x fixed, is $y \mapsto f(x,y)$ a measurable function, and is $F(x) := \int f(x,y)dy$ a measurable function? Regarding the first question, let's consider $f(x,y) = \chi_E(x,y),\ E \subseteq \mathbb{R}^{n+m}$ measurable. Write $f(x,\cdot)$ as a function of y where x is fixed. Thus, we have

$$f(x,y) = \begin{cases} 1 \text{ if } y \in E_x \\ 0 \text{ if } y \notin E_x. \end{cases}$$

We then need to think about what E_x is. For a fixed x, $E_x = \{y \in \mathbb{R}^m\ :\ (x,y) \in E\}$. In other words, these are the slices of E. So $f(x,\cdot) = \chi_{E_x}$ is measurable if and only if E_x measurable for fixed x. If E is measurable, is it always true that E_x must be measurable? The answer to this question turns out to "almost."

Example 4.3. Take $F \subseteq [0,1]$ to be a non-measurable set. Take

$$E = [(0,1) \times (0,1)] \cup [\{1\} \times F].$$

Then we see that $E_1 = F$, a non-measurable set, even though we have that E is measurable (the slice where it is not measurable has measure 0). This provides us with the key – we'll be able to say that E_x is measurable a.e.

We also see that the conclusion of Fubini is not always true.

Example 4.4. We define a function f over $\mathbb{R}^3$, which will be hard to graph. Instead, we will just describe how it looks. Form a sequence of unit squares, starting from the origin and moving towards infinity diagonally. Divide the squares into fourths, and in each square repeat the same pattern. That is, denoting the nth unit square by S_n, we have $x \in S_1$, $y \in S_1$,

$$f(x,y) = \begin{cases} 1 \text{ if } 0 \leqslant x,y \leqslant 1/2 \text{ or } 1/2 \leqslant x,y \leqslant 1, \\ 0 \text{ if } 0 \leqslant x \leqslant 1/2, 1/2 \leqslant y \leqslant 1 \text{ or } 0 \leqslant y \leqslant 1/2, 1/2 \leqslant x \leqslant 1. \end{cases}$$

Repeat this pattern for all n. Then we see that

$$\int\int |f| = \int f^+ + \int f^- = \infty + \infty = \infty,$$

so f is not integrable. Furthermore, $\int\int f$ is not defined, since we have $\infty - \infty$. On the other hand, $\int f(x,y)dy = 0$ for any x.

Let us define a class of functions which will be helpful moving forward.

Definition. Let $\mathcal{F}$ be the class of measurable functions, f, on $\mathbb{R}^{n+m}$ such that

(i) $f \in L(\mathbb{R}^{n+m})$,

(ii) For almost every x, $f(x,\cdot)$ is measurable and integrable on $\mathbb{R}^m$,

(iii) $F(x) := \int f(x,y)dy$ is measurable and integrable on $\mathbb{R}^n$ and $\int\int f = \int F$.

This leads us to Fubini's theorem.

Theorem 4.16 (Fubini's Theorem)**.** If $f \in L(\mathbb{R}^{n+m})$, then $f \in \mathcal{F}$.

We will prove this using a series of lemmas. We will prove the following lemma in this lecture, and will finish the proof in the next.

Lemma 4.5. We have that $\mathcal{F}$ is closed under linear combinations and under monotone limits, as long as the limit is integrable.

Proof. We first prove it for linear combinations. Take $f = \sum_{i=1}^N a_i f_i$, where $f_i \in \mathcal{F}$. Let $Z_i = \{x \ L \ f_i(x,\cdot) \notin \mathcal{M}\}$. By assumption, $|Z_i| = 0$. Let $Z = \bigcup_{i=1}^N Z_i$. We note that $|Z| = 0$. So for $x \notin Z$, we get $f_i(x,\cdot)$ is measurable and also integrable. We therefore know that $f = \sum_{i=1}^N a_i f_i(x,\cdot)$ is measurable and integrable. Define $F(x) := \int \sum_{i=1}^N a_i f_i(x,\cdot)$ if $x \notin Z$ and 0 otherwise (it's fine to do this since Z has measure 0). Then using linearity, we get

$$F(x) = \int \sum_{i=1}^N a_i f_i(x,\cdot) = \sum_{i=1}^N a_i \int f_i(x,\cdot) = \sum_{i=1}^N a_i F_i(x).$$

Furthermore, integrating $F(x)$ gives us

$$\int F(x) = \int \sum_{i=1}^N a_i F_i(x) = \sum_{i=1}^N a_i \int F_i(x) = \int\int f,$$

and so we win.

We now prove it holds for monotone limits. Assume that $f_k \nearrow f \in L(\mathbb{R}^{n+m})$, $f_k \in \mathcal{F}$. Let $Z_k = \{x \ : \ f_k(x,\cdot) \notin \mathcal{M}\}$. Set $Z = \bigcup_k Z_k$, and note $|Z| = 0$. For $x \notin Z$, we have $f_k(x,\cdot) \nearrow f(x,\cdot)$. Furthermore, since this is an increasing sequence, we have $f_1(x,\cdot) \leqslant f_k(x,\cdot) \leqslant f(x,\cdot)$ for all k, and so we get

$$\int f_k(x,\cdot)dy \to \int f(x,\cdot)dy$$

by the Monotone Convergence Theorem (**Theorem 4.8**). We then get that the integral is well-defined, and exists, but we do not know that $f(x,\dot{)}$ is integrable, since it might infinite. Let $F_k(x0 = \int f_k(x,y)dy$, and again let $F_k(x) = 0$ if $x \in Z$. Then we have

$$F_k(x) \to F(x).$$

Furthermore, we have $F_k(x) \nearrow F(x)$. Thus, $F(x)$ is measurable. We use the MCT again to get

$$\int F_k \to \int F,$$

noting again that $F_1(x) \leqslant F_k(x)$. Now

$$\int \int f_k \to \int \int f$$

as $k \to \infty$ by MCT. Furthermore, we have

$$\int \int f_k = \int F_k,$$

so this implies

$$\int F = \int \int f.$$

Since $f \in L(\mathbb{R}^{n+m})$, we have that $\int F < \infty$ and integrable, which forces $F < \infty$. This then implies that $f(x,\cdot)$ is integrable as well. This gives us the result. **Q.E.D**

4.6 Lecture 16 (Proving Fubini, Tonelli's Theorem)

Lemma 4.6. If $E \subseteq \mathbb{R}^{n+m}$ is measurable and $|E| < \infty$, then $\chi_E \in \mathcal{F}$.

Proof. We prove this by building things up.

(i) We start with half open intervals. That is, the sets

$$J = [a_1, b_1) \times \cdots \times [a_{n+m}, b_{n+m}).$$

Notice that we may write this as $J = J_1 \times J_2$, where

$$J_1 = [a_1, b_1) \times \cdots \times [a_n, b_n)$$

and

$$J_2 = [a_{n+1}, b_{n+1}) \times \cdots \times [a_{n+m}, b_{n+m}).$$

Then we may write $\chi_J(x,y) = \chi_{J_1}(x)\chi_{J_2}(y)$. Now, we want to prove that χ_J is in $\mathcal{F}$. We first note that $\chi_J \in L(\mathbb{R}^{n+m})$ trivially, since it is a characteristic function over a set of finite measure. Next, we'd like to show that for almost all x, $\chi_J(x,\cdot)$ is measurable and integrable. Since $\chi_J(x,\cdot) = \chi_{J_1}(x)\chi_{J_2}(\cdot)$, we have that $\chi_J(x,\cdot) = \chi_{J_2}(\cdot)$ or $\chi_J(x,\cdot) = 0$, depending on whether $x \in J_1$ or not. So it is measurable for all x, and furthermore it is integrable since it is either a characteristic function or zero. Now, we need to show that

$$F(x) := \int \chi_J(x,y)dy$$

is measurable and integrable as well. We rewrite this as

$$F(x) = \int \chi_{J_1}(x)\chi_{J_2}(y)dy = \chi_{J_1}(x)\int \chi_{J_2}(y)dy = \chi_{J_1}(x)|J_2|.$$

Since $|J_2| < \infty$, we have that this is a measurable function. Furthermore, integrating this over x gives us

$$\int F(x)dx = |J_2|\int \chi_{J_1}(x)dx = |J_1||J_2|.$$

We know from prior that

$$|J_1||J_2| = |J_1 \times J_2| = |J| = \int\int \chi_J(x,y).$$

So we have that it is integrable, and that it is equal to the double integral. This gives us that it must be in $\mathcal{F}$.

(ii) Now let's prove it for open sets. Let $G \subseteq \mathbb{R}^{n+m}$ be open. Then we know that we may cover G by non-overlapping closed intervals. Taking these intervals to be half open, we get that they must actually be disjoint. Hence, we may write

$$G = \bigsqcup_{k=1}^{\infty} J_k.$$

Now, taking the characteristic functions, we have

$$\chi_G = \sum_{k=1}^{\infty} \chi_{J_k}.$$

Let's cap this off at n; that is, examine

$$f_n = \sum_{k=1}^{n} \chi_{J_k}.$$

Since $\chi_{J_k} \in \mathcal{F}$ for each k, **Lemma 4.5** tells us that the sum from 1 to n of these functions must also be in $\mathcal{F}$. Now, notice that $f_n \nearrow f = \sum_{k=1}^{\infty} \chi_{J_k}$. **Lemma 4.5** tells us that $\mathcal{F}$ is closed under monotone limits, and so we get that $f \in \mathcal{F}$ as well. Therefore, we have that $\chi_G = f$ is in $\mathcal{F}$.

(iii) Since we have it for open sets, we can also show it holds for sets of type G_δ. Let H be such a set. Then we have $H = \bigcap_k G_k$, where G_k are open for all k. Without loss of generality, take these to be decreasing sets. Then we have that $\chi_{G_k} \searrow \chi_H$. Again, this is a monotone limit, and so since $\chi_{G_k} \in \mathcal{F}$ we must have $\chi_H \in \mathcal{F}$.

(iv) We also need to show it holds for sets of measure 0. Let Z be such a set. Then we can find a G_δ set H such that $Z \subseteq H$ and $|H| = |Z| = 0$. We must now show $\chi_Z \in \mathcal{F}$. We have that $\chi_Z \in L(\mathbb{R}^{n+m})$ by default. Fixing x, we have that $\chi_Z(x, \cdot)$ is measurable a.e. since $\chi_Z(x, \cdot) = \chi_H(x, \cdot) = 0$ for a.e. x. Furthemore, we see that $\chi_Z(x, \cdot)$ is integrable. Let

$$F(x) = \int \chi_Z(x, y)dy = 0.$$

We get that this is therefore measurable and integrable. Furthermore,

$$\int F(x)dy = \int\int \chi_Z dxdy = 0.$$

(v) For general measurable sets E, **Theorem 2.3 (ii)** says that we may write it as $E = H \backslash Z$, where H is of type G_δ and $|Z| = 0$. Hence, $\chi_E = \chi_H - \chi_Z$. Since $\chi_H, \chi_Z \in \mathcal{F}$, we get that $\chi_E \in \mathcal{F}$ by using the fact that it's closed under linear combinations.

Q.E.D

We now finally prove Fubini's theorem.

Proof. We again build things up.

(i) Simple functions which are integrable are in $\mathcal{F}$ by **Lemma 4.5** and **Lemma 4.6**.

(ii) Non-negative integrable functions f are in $\mathcal{F}$, since **Theorem 3.9** gives us that we have a sequence of simple measurable functions $f_k \nearrow f$, and $\mathcal{F}$ is closed under monotone limits.

(iii) For general f, we write it as $f = f^+ - f^-$. Since each of these are in $\mathcal{F}$, and $\mathcal{F}$ is closed under linear combinations, we have that $f \in \mathcal{F}$.

Q.E.D

We also have a similar theorem called Tonelli's theorem, which is a very close relative to Fubini's theorem.

Theorem 4.17 (Tonelli's Theorem)**.** If $f(x, y)$ is a non-negative measurable function on $\mathbb{R}^{n+m}$, then

$$\int\int f dxdy = \int \left(\int f dx \right) dy = \int \left(\int f dy \right) dx,$$

and $f(x, \cdot)$ and $F(x) = \int f(x, y)dy$ are measurable.

Proof. It turns out that this is just a consequence of Fubini's theorem. Define $a \wedge b = \min\{a, b\}$. Let

$$f_n = (f(x,y) \wedge n)\chi_{\{(x,y)\ :\ |(x,y)|<n\}}$$

. Then we see that f_n is just a linear multiple of a characteristic function, and so $f_n \in \mathcal{F}$ for all n. Furthermore, let $Z_n = \{x \ :\ f_n(x,\cdot) \notin \mathcal{M}\}$. Then $|Z_n| = 0$. Letting $Z = \bigcup_n Z_n$, we have $|Z| = 0$. Furthermore, if $x \notin Z$, $f(x,\cdot) = \lim_{n\to\infty} f_n(x,\cdot)$, and is therefore measurable. Similarly, the MCT tells us

$$\int f_n(x,y)dy \to \int f(x,y)dy.$$

This tells us that $F_n(x) \to F(x)$, and that they are measurable and integrable. Finally, we see that

$$\int F = \lim_{n\to\infty} \int F_n = \lim_{n\to\infty} \int\int f_n = \int\int f.$$

Thus, $f \in \mathcal{F}$. **Q.E.D**

We now go on to discuss some applications.

Lemma 4.7. If $\sum_{k=1}^{\infty} |f_k| \in L(\mathbb{R}^n)$, then $\int \sum_{k=1}^{\infty} f_k = \sum_{k=1}^{\infty} \int f_k$.

Proof. This will follow later by seeing that sums are just discrete integrals. For now, let

$$f(x,y) = \sum_{k=1}^{\infty} f_k(x)\chi_{[k,k+1)}(y).$$

Then we see that

$$\int \left(\int f(x,y)dy \right) dx = \int \sum_{k=1}^{\infty} f_k(x)dx.$$

Q.E.D

Lemma 4.8. We have that

$$z^p = \int_0^z py^{p-1}dy.$$

Lemma 4.9. For any f, we have

$$\int |f|^p dx = \int_0^{\infty} py^{p-1}\omega_{|f|}(y)dy.$$

Proof. Using the prior lemma, we have

$$\int |f|^p dx = \int \left(\int_0^{|f|} py^{p-1}dy \right) dx = \int \left(\int_0^{\infty} py^{p-1}\chi_{y\leqslant|f|}dy \right) dx$$

$$= \int_0^\infty \left(\int p y^{p-1} \chi_{y \leqslant |f|} dx \right) dy = \int_0^\infty p y^{p-1} \left(\int \chi_{y \leqslant |f|} dx \right) dy.$$

Solving the integral

$$\int \chi_{y \leqslant |f|} dx,$$

we see that we just ge the measure of the set $\{x \ : \ |f| \geqslant y\}$, which we may rewrite as $\omega_{|f|}(y) = |\{x \ : \ |f(x)| \geqslant y\}|$ (the reason we may drop the inequality is that $\omega_{|f|}$ is a non-increasing function, and so has finitely many discontinuities). Putting this together, we get

$$\int |f|^p dx = \int_0^\infty p y^{p-1} \omega_{|f|}(y) dy,$$

an improper Riemann–Stieltjes integral. **Q.E.D**

4.7 Lecture 17 (Convolutions)

We finish this section by talking a little about convolutions, which gives us a final application of Fubini's theorem.

Definition. If f, g are measurable functions on $\mathbb{R}^n$, then the **convolution** $f * g$ is defined by

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x-y) g(y) dy,$$

so long as the integral exists.

Remark 20. One quick remark is that we have $f * g = g * f$ through a change of variables.

Theorem 4.18. If $f, g \in L(\mathbb{R}^n)$, then $f * g$ exists a.e. is in $L(\mathbb{R}^n)$, and

$$\int |f * g| \leqslant \left(\int |f| \right) \left(\int |g| \right.$$

Furthermore, if $f, g \geqslant 0$, then

$$\int f * g = \left(\int f \right) \left(\int g \right).$$

Before proving the theorem, we need a lemma.

Lemma 4.10. If f is measurable on $\mathbb{R}^n$, then $F(x, y) := f(x-y)$ is measurable on $\mathbb{R}^{2n}$.

Proof. The proof is really a geometric proof. We want to show that $\{(x, y) \in \mathbb{R}^{2n} \ : \ F(x, y) > a\}$ is a measurable set for all a. But this, in particular, is just the set $\{(x, y) \in \mathbb{R}^{2n} \ : \ f(x-y) > a\}$ for all a. Notice $F(x_0, y_0) > a$ implies that the hyperplane $x - b = y$, where $b = x_0 - y_0$ also satisfies this property. Since the slope is given, we have that these lines are completely determined by their y intercept. So, we may look at $\{s \in \mathbb{R}^n : f(s) > a\} \times \mathbb{R}^n$, which is a measurable set. Using a Lipschitz transformation T which maps this set to $\{(x, y) \in \mathbb{R}^{2n}\}$ by multiplying the values together tells us that $\{(x, y) : F(x, y) > a\}$ is measurable. **Q.E.D**

Proof. We start by proving the second half. Let $f, g \geqslant 0$. Then

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

by Tonelli's theorem. Now, since y is fixed, we use the fact that the Lebesgue integral for non-negative functions is simply a measure of area underneath the curve to see that

$$\int f(x-y)dx = \int f(x)dx.$$

Hence, we can pull everything out to get

$$\int (f * g) = \left(\int f \right) \left(\int g \right).$$

Now, notice that

$$|f * g| = \left| \int f(x-y)g(y)dy \right| \leqslant \int |f(x-y)||g(y)dy = |f| * |g|,$$

using the triangle inequality. Hence, we have

$$|f * g| \leqslant \left(\int |f| \right) \left(\int |g| \right).$$

Since $f, g \in L(\mathbb{R}^n)$, then so is $|f|, |g|$. In particular, this tells us that $f * g \in L(\mathbb{R}^n)$. **Q.E.D**

Chapter 5

Lebesgue Differentiation

5.1 Lecture 18 (Lebesgue Differentiation)

We now want to find an analogue for the F.T.C. for Lebesgue integrals. One important consequence of this is that we will develop a characterization of when

$$f(b) - f(a) = \int_a^b f'(x)dx$$

holds.

Definition. For an integrable function f on $\mathbb{R}^n$, the **indefinite integral** F is defined by

$$F(A) = \int_A f(x)dx.$$

Notice that A is a set here.

Definition. A **set function** is a function on a σ-algebra Σ that is finite for all $A \in \Sigma$ and is countable additive; that is, if $\{A_i\}$ are disjoint in $\Sigma <$ then

$$F\left(\bigcup A_n\right) = \sum F(A_n).$$

Remark 21. We see that the indefinite integral F satisfies being a set function.

Definition. For $f \in L(\mathbb{R}^n)$, the indefinite integral F is **differentiable** at $x \in \mathbb{R}^n$ with derivative $f(x)$ if

$$\lim_{r\to 0} \frac{1}{|Q_r(x)|} F(Q_r(x)) = f(x),$$

where $Q_r(x)$ is the cube centered at x with sidelength $2r$.

We state the following theorem, though it will take a few lectures to prove.

Theorem 5.1. (Lebesgue Differentiation Theorem) If $f \in L(\mathbb{R}^n)$, then the indefinite integral F is differentiable, with derivative $f(x)$ at almost every x.

Remark 22. (1) In dimension 1, this says

$$\lim_{r\to 0} \frac{1}{2r} \int_{x-r}^{x+r} f(y)dy = f(x)$$

a.e. We'll see later that it's also true

$$\lim_{r\to 0} \frac{1}{r} \int_{x}^{x+r} f(y)dy = f(x).$$

That is, the average property over a very small interval must just be the value at the (center) of the inteval.

(2) This is (obviously) true for continuous functions.

First, we need to see why it is true that $\lim_{r\to 0} F(Q_r(x)) = 0$. To see this, notice that we have

$$\lim_{r\to 0} F(Q_r(x)) = \lim_{r\to 0} \int_{Q(x)} f(y)dy$$

by definition. Now, we can switch this to be an integral over all of $\mathbb{R}^n$ by using a characteristic function. That is,

$$\lim_{r\to 0} \int_{Q(x)} f(y)dy = \lim_{r\to 0} \int_{\mathbb{R}^n} f(y)\chi_{Q_r(y)}dy.$$

Now, we want to be able to justify pulling the limit into the integral. One way to do so is by using the Dominated Convergence Theorem. That is, we have that this is dominated by $f(x)$, and we know $f(x)$ satisfies the conditions. So pulling the limit in, we get

$$\lim_{r\to 0} \int_{\mathbb{R}^n} f(y)\chi_{Q_r(y)}dy = \int_{\mathbb{R}^n} f(y) \lim_{r\to 0} \chi_{Q_r(y)}dy = 0.$$

We now talk about a few more definitions.

Definition. An abstract set function F is **continuous** if

$$\lim_{\operatorname{diam}(A)\to 0} F(A) = 0.$$

What we've just established is that F is continuous. But we can get even stronger results.

Definition. A set function F is **absolutely continuous** if

$$\lim_{|A|\to 0} F(A) = 0.$$

We now want to show that F is absolutely continuous.

Theorem 5.2. If $f \in L(\mathbb{R}^N)$, then F is absolutely continuous.

Proof. Let $\epsilon > 0$, then we want to find a $\delta > 0$ so that $|A| < \delta$ implies $|F(A)| < \epsilon$. First, we note that

$$\int_{\{f \geqslant n\}} f \to 0$$

as $n \to \infty$. So we can take n sufficiently large so that

$$\left| \int_{\{f \geqslant n\}} f \right| < \frac{\epsilon}{2}.$$

Now,

$$|F(A)| \leqslant \left| \int_{A \cap \{f \geqslant n\}} f \right| + \left| \int_{A \cap \{f < n\}} f \right|.$$

We know that we can bound this by

$$\left| \int_{A \cap \{f \geqslant n\}} f \right| + \left| \int_{A \cap \{f < n\}} f \right| < \frac{\epsilon}{2} + \left| \int_{A \cap \{A \cap \{f < n\}} f \right|.$$

Since we have $|A| < \delta$, and f bounded above by n, take $|A| < \epsilon/(2n)$. Then we get

$$|F(A)| < \epsilon,$$

as desired. **Q.E.D**

5.2 Lecture 19 (Proving Lebesgue Differentiation)

To prove the theorem, we're going to set up a few lemmas ahead of time and then work backwards proving those lemmas.

Lemma 5.1. If $f \in L(\mathbb{R}^n)$, then there exists a sequence ϕ_k of continuous functions with compact support such that

$$\lim_{k \to \infty} \int |f - \phi_k| = 0.$$

We say that this converges in L^1.

Remark 23. We had before that the Lebesgue differentiation theorem clearly held true for continuous functions, so this gets us somewhat close to showing it for measurable functions.

Definition. If $f \in L(\mathbb{R}^n)$, we define the **Hardy-Littlewood maximal function** f^* by

$$f^*(x) := \sup_{r>0} \frac{1}{|Q_r(x)|} \int_{Q_r(x)} |f|.$$

The Hardy-Littlewood function is asking what is the biggest average value that we could get. This will help us establish whether the limit makes sense in the first place. Using this, we establish the second lemma we'll need (which seems a little like Chebychev's).

Lemma 5.2. If $f \in L(\mathbb{R}^n)$, then there exists a constant depending only on n, C_n, such that

$$|\{f^* > a\}| \leqslant \frac{C_n}{a} \int |f|.$$

Note that as $a \to \infty$, the right hand side goes to 0, so we get that f^* is finite almost everywhere. In other words, we have that f^* is **almost integrable**; an idea that we'll formalize. Note that if $g \in L(\mathbb{R}^n)$, then

$$|\{g > a\}| \leqslant \frac{1}{a} \int |g|.$$

Definition. If we have

$$|\{g > a\}| \leqslant \frac{c}{a},$$

then we say that g is **weakly integrable**, and we say that is in **weak-$L(\mathbb{R}^n)$**.

Remark 24. We have that the function f^* is **never** integrable unless $f = 0$ identically.

We now prove the Lebesgue Differentiation Theorem using these tools.

Proof. For notation sake, let ϕ be a continuous function with compact support and let $Q = Q_r(x)$. We then want to compare

$$\left| \frac{1}{|Q|} \int_Q f(y)dy - f(x) \right| \leqslant \left| \frac{1}{|Q|} \right|.$$

Notice that we can add and subtract terms as well as use the triangle inequality to get the following upper bound;

$$\left| \frac{1}{|Q|} \int_Q f - \frac{1}{|Q|} \int_Q \phi \right| + \left| \frac{1}{|Q|} \int_Q \phi - \phi(x) \right| + |\phi(x) - f(x)|.$$

Now, as $Q \searrow x$, we have that

$$\left| \frac{1}{|Q|} \int_Q \phi - \phi(x) \right| \to 0.$$

So we can say (after bringing the absolute value in)

$$\limsup_{r\to 0}\left|\frac{1}{|Q|}\int_Q f - f(x)\right| \leqslant \limsup_{r\to 0}\frac{1}{|Q|}\int_Q |f-\phi| + |\phi(x)-f(x)|.$$

We can then bound this above by

$$(f-\phi)^*(x) + |\phi(x)-f(x)|.$$

So looking at the sets, we have

$$\left|\left\{x \ : \ \limsup_{r\to 0}\left|\frac{1}{|Q|}\int_Q f - f(x)\right| > \epsilon\right\}\right| \leqslant \left|\left\{(f-\phi)^* > \frac{\epsilon}{2}\right\}\right| + \left|\left\{|\phi - f| > \frac{\epsilon}{2}\right\}\right|.$$

We use Chebychev to get

$$|\{|\phi - f| > \epsilon/2\}| \leqslant \frac{2}{\epsilon}\int |\phi - f|,$$

and we use **Lemma 5.2** to get

$$|\{(f-\phi)^* > \epsilon/2\}| \leqslant \frac{2C_n}{\epsilon}\int |\phi - f|.$$

So we can put things together based on these. Nothing on the left depends on ϕ, so we can choose ϕ based on **Lemma 5.1** so that everything on the right goes to 0, thus giving us our desired result. **Q.E.D**

We now prove **Lemma 5.1**.

Proof. Let $\mathcal{A}$ be the class of functions such that the lemma holds. We will show that $L(\mathbb{R}^n) \subseteq \mathcal{A}$. We do this by working our way up.

Claim 5.1. $\mathcal{A}$ is closed under linear combinations.

Proof. Take $\sum a_i f_i$ where $f_i \in \mathcal{A}$. Then for each f_i we can create a sequence $\phi_{i,k} \to f_i$. We guess then that the appropriate sequence is $\sum a_i \phi_{i,k}$. To show that this is the correct sequence, we check

$$\int \left|\sum_i a_i f_i - \sum_i a_i \phi_{i,k}\right| \leqslant \sum_i |a_i| \int |\phi_{i,k} - f_i| \to 0.$$

Q.E.D

Claim 5.2. $\mathcal{A}$ is closed in L^1; that is, if $f_k \in \mathcal{A}$, and

$$\int |f_k - f| \to 0,$$

then $f \in \mathcal{A}$.

Proof. Take $\phi_k \to f_k$. Then we have

$$\int |f - \phi_k| \leqslant \int |f - f_k| + \int |f_k - \phi_k| \to 0.$$

Q.E.D

We now build up sets. It's clear that for any interval $I \subseteq \mathbb{R}^n$, we have that $\chi_I \in \mathcal{A}$. Next, let G be open and $|G| < \infty$. Then we have that there are a sequence of non-overlapping intervals I_n so that

$$G \subseteq \bigcup_n I_n.$$

Since they are non-overlapping, we get

$$\chi_G = \sum_n \chi_{I_n} \text{ a.e.}$$

Now, take the finite partial sums. Then we have

$$\sum_{n=1}^{N} \chi_{I_n} \nearrow \chi_G.$$

Notice that for each N, $\sum_{n=1}^{N} \chi_{I_n} \in \mathcal{A}$. We then check

$$\int \left| \chi_G - \sum_n \chi_{I_n} \right| = \int \left(\chi_G - \sum_n \chi_{I_n} \right) \to 0$$

by the MCT.

Now, we want to get general measurable sets E with $|E| < \infty$. We know $E = H \backslash Z$, $|Z| = 0$ and H of type G_δ. So there exists a sequence of open sets $G_k \searrow H$, and so $\chi_{G_k} \searrow \chi_E$ a.e. Therefore, we have

$$\int |\chi_{G_k} - \chi_E| \to 0$$

again by the MCT. So we thus get all measurable functions. **Q.E.D**

To prove **Lemma 5.2**, we will need another lemma.

Lemma 5.3. (Simple Vitali Lemma) Let $E \subseteq \mathbb{R}^n$ with $|E|_e < \infty$ and let K be a collection of cubes that cover E. Then there exists a finite collection of disjoint cubes $\{Q_k\}_{k=1}^{n} \subseteq K$ such that

$$\sum_{k=1}^{N} |Q_k| \geqslant \frac{1}{2 \cdot 5^n} |E|_e.$$

We will prove this lemma in the next lecture. For now, we use it to prove **Lemma 5.2**.

Proof. For $R < \infty$, let $E_k = \{f^* > a\} \cap \{|x| < R\}$. If $x \in E_k$, then there exists a cube Q_x such that

$$\frac{1}{|Q_x|}\int_{Q_x}|f| > a \iff |Q_x| < \frac{1}{a}\int_{Q_x}|f|.$$

Let $K = \{Q : Q \text{ satisfies the inequality above}\}$. Note that $E_R \subseteq \bigcup_{Q\in K} Q$. We can use the Simple Vitali Lemma to find $Q_1, \ldots, Q_N$ such that

$$\sum_{k=1}^{N}|Q_k| > \frac{1}{2\cdot 5^n}|E_k|.$$

On the other hand, $|E_R| < (2\cdot 5^n)\sum_{k=1}^{N} Q_K$. But the measure of those cubes satisfies

$$|Q_k| < \frac{1}{a}\int_{Q_x}|f|.$$

So therefore we have

$$|E_R| < \frac{2\cdot 5^n}{a}\sum\int_{Q_x}|f| \leqslant \frac{2\cdot 5^n}{a}\int|f|,$$

since the cubes were disjoint. Finally, $E_R \nearrow \{f^* > a\}$, and so

$$|\{f^* > a\}| < \frac{2\cdot 5^n}{a}\int|f|.$$

Q.E.D

5.3 Lecture 20 (Proving Simple Vitali Lemma)

We'll first note that we can actually say something stronger about the Lebesgue differentiation theorem, which is that it holds if f is just **locally integrable**, instead of globally integrable. We formalize this thought now.

Definition. A function f is **locally integrable** if it is integrable on all bounded subsets. We denote this by $f \in L_{\text{loc}}(\mathbb{R}^n)$.

It's clear that theorem holds under only the assumption of locally integrable, since it is really a local problem. We can also make yet another stronger statement. We first need to declare what a Lebesgue point of f is.

Definition. We say that x is a **Lebesgue point of f** if

$$\lim_{r\to 0}\frac{1}{|Q_r(x)|}\int_{Q_r(x)}|f(y) - f(x)|dy = 0.$$

The absolute value is what makes this stronger.

Theorem 5.3. If f is locally integrable, then almost every $x \in \mathbb{R}^n$ is a Lebesgue point of f.

Proof. For every $q \in \mathbb{R}$, we have that $|f - q|$ is still locally integrable. Now, by assumption note that every point in $\mathbb{R}^n$ is a Lebesgue point for $|f - q|$, where we take $q \in \mathbb{Q}$. Then we have

$$\frac{1}{|Q|}\int_Q |f(y) - f(x)|dy \leqslant \frac{1}{|Q|}\int_Q |f(y) - q|dy + |q - f(x)|.$$

Letting $|Q| \to 0$, we see that

$$\frac{1}{|Q|}\int_Q |f(y) - q|dy \to |f(x) - q|.$$

So we have

$$\limsup_{r\to 0} \frac{1}{|Q_r(x)|}\int_{Q_r(x)} |f(y) - f(x)|dy \leqslant 2|q - f(x)|.$$

Since $\mathbb{Q}$ is dense, we can take q arbitrarily close to $f(x)$, giving us that this goes to 0. **Q.E.D**

We also note that we don't have to just take cubes, but can take sets which satisfy some strict properties.

Definition. A family of sets S is **regularly shrinking to a point x** if

(i) $\operatorname{diam}(S) \to 0$,

(ii) There is a constant k such that if Q is the smallest cube centered at x containing S, then

$$|Q| < k|S|.$$

Example 5.1. The set $[x, x+h]$ is regularly shrinking, since the smallest cubes containing this are $[x-h, x+h]$ and we have

$$|Q| \leqslant 2|S|.$$

Theorem 5.4. If $f \in L_{\text{loc}}(\mathbb{R}^n)$, then at every Lebesgue point,

$$\lim_{S\searrow\{x\}} \frac{1}{|S|}\int_S |f(y) - f(x)|dy = 0$$

for any family of sets S that is regularly shrinking to x.

Proof. We have

$$\frac{1}{|S|}\int_S |f(y) - f(x)|dy \leqslant \frac{1}{|S|}\int_Q |f(y) - f(x)|dy,$$

where $S \subseteq Q$. We can then bound this above by

$$\frac{k}{|Q|}\int_Q |f(y) - f(x)|dy \to 0.$$

Q.E.D

We implicitly used the fact that f^* is a measurable function. We now prove this.

Claim 5.3. The Hardy-Littlewood maximal function is a measurable function. More specifically, the set $\{f^* > a\}$ is an open set for all a.

Proof. Select $x \in \{f^* > a\}$. Then we want to show that every point which is sufficiently close to x is also greater than a. Assume for contradiction that this is not the case. That is, we have a sequence $x_k \to x$ where $f^*(x_k) \leqslant a$ for all k but $f^*(x) > a$. Using the definition, we see that this is saying

$$\frac{1}{|Q_{r'}(x_k)|}\int_{Q_{r'}(x_k)} f(y)dy,$$

where we choose r' such that the cube $Q_{r'}(x_k)$ is the biggest cube contained in $Q_r(x)$. The dominated convergence theorem tells us that

$$\frac{1}{|Q_{r'}(x_k)|}\int_{Q_{r'}(x_k)} f(y)dy \to \frac{1}{|Q_r(x)|}\int_{Q_r(x)} f(y)dy,$$

but this gives us a contradiction, since this would force the right hand side to be less than or equal to a. **Q.E.D**

Remark 25. This, in fact, proves that f^* is lsc.

We start the proof of the Simple Vitali Lemma (**Lemma 5.3**).

Proof. Without loss of generality, assume that the cubes are not arbitrarily large (if they were, we win by default). Let d_1 be the supremum of the side lengths of cubes in K. Pick Q_1 with side length greater than or equal to $d_1/2$. Now, decompose K into two parts, denoted by K_2 and K_2', where K_2' is the collection of cubes intersecting Q_1 and K_2 is the collection of cubes disjoint from Q_1. Let d_2 be the supremum of the side lengths of cubes in K_2. Choose Q_2 so that it has side length greater than or equal to $d_2/2$. Repeat this process. Note that we have formed a non-increasing sequence $d_1 \geqslant d_2 \geqslant \cdots$. We then divide this up into cases.

Case 1 If the $d_j \geqslant \delta$ for all j, $\delta > 0$, then we have an infinite sequence of cubes that are disjoint, and they all have side length greater than or equal to $\delta/2$. This then says $|Q_j| \geqslant (\delta/2)^n$ for all J. We just have to pick N large enough, then.

Case 2 What if $d_N > 0$, $d_{N+1} = 0$. Notice then that we have

$$K = \bigcup_{n=2}^{N} K_n'.$$

Let $Q_j \in K_{j+1}'$, and pick $Q \in K_{j+1}'$. For any such Q, we claim $Q \subseteq \widetilde{Q_j}$, which is the cocentric cube centered at Q_j but with 5 times the edge length. This is becasue the side length of the Q's are at most 2 times the side length of Q_j. Then, since K covers E, we have that the set E is covered by these blown up cubes. Thus, we get

$$|E|_e \leqslant \sum_{j=1}^{N} |\widetilde{Q_j}| = \sum_{j=1}^{N} 5^n |Q_j|,$$

which tells us

$$\sum_{j=1}^{N} |Q_j| > \frac{1}{5^n}|E|_e > \frac{1}{2 \cdot 5^n}|E|_e.$$

5.4 Lecture 21 (Vitali Covering Lemma)

Case 3 Now consider the case where $d_n \to 0$, $d_n > 0$ for all n. Then we may have an infinite sequence Q_j of disjoint cubes, and we denote $\widetilde{Q_j}$ to be the blow up from Case 2. Recall that if a cube $Q \in K'_{j+1}$, then $Q \subseteq \widetilde{Q_j}$. Now, we use this to claim that if $Q \in K$, then Q is covered by $\bigcup_j \widetilde{Q_j}$. Assume that is not; that is, $Q \nsubseteq \bigcup_j \widetilde{Q_j}$. Then $Q \notin K'_{j+1}$ for every j, so $Q \in K_{j+1}$ for every j. This means that the side length of $Q \leqslant d_{j+1}$ for all j. Since $d_n \to 0$, we must have that the side length of Q goes to 0 as well.

Remark 26. Here, we note that cubes are not points. That is, a cube can not have side length 0.

So we have that $E \subseteq \bigcup_j \widetilde{Q_j}$. Then

$$|E|_e \leqslant \sum_j |\widetilde{Q}_j| \leqslant 5^n \sum_j Q_j.$$

Thus, we have

$$\sum_{j=1}^{\infty} |Q_j| \geqslant \frac{1}{5^n}|E|_e.$$

Now, we pick N large enough so that

$$\sum_{j=1}^{N} |Q_j| \geqslant \frac{1}{2 \cdot 5^n}|E|_e.$$

Q.E.D

We now want to cover the Big Vitali Lemma, or the Vitali Covering Lemma. First, we need a definition.

Definition. A collection K covers E in the **Vitali sense** if, for every $x \in E$ and every $\delta > 0$, there is a cube in K containing x with diameter less than δ.

This will be helpful in generalizing Vitali's lemma.

Theorem 5.5. (Vitali Covering Lemma) If $E \subseteq \mathbb{R}^n$ is covered in the Vitali sense by K, and E is a set such that $0 < |E|_e < \infty$, then given $\epsilon > 0$ there exists a sequence $\{Q_j\}_{j=1}^{\infty}$ of disjoint cubes with the following two properties:

(i)
$$\left|E - \bigcup_{j=1}^{\infty} Q_j\right| = 0$$

(ii)
$$\sum_j |Q_j| < (1+\epsilon)|E|_e.$$

Proof. First, pick an open set G such that $E \subseteq G$ and $|G| < (1+\epsilon)|E|_e$. Without loss of generality, assume all cubes in K are contained in G, since the cubes cover in the Vitali sense. So al lwe need to do is get the first property, since the second property comes from the fact that

$$\left|\bigcup_j Q_j\right| = \sum_j |Q|_j \leqslant |G| < (1+\epsilon)|E|_e.$$

Without loss of generality, take $\epsilon < (3/10)(1/5^n)$. If we show this for such ϵ, then we can get that it holds for larger ϵ by just capping it off.

By the Simple Vitali Lemma (**Lemma 5.3**), there exists a sequence of disjoint cubes $\{Q_j\}_{j=1}^{N_1}$ with

$$\sum_{j=1}^{N_1} |Q_j| \geqslant \frac{1}{2 \cdot 5^n}|E|_e.$$

Then notice
$$\left|E - \bigcup_{j=1}^{N_1} Q_j\right| \leqslant \left|G - \bigcup_{j=1}^{N_1} Q_j\right| = |G| - \sum_{j=1}^{N_1} |Q_j|$$

$$\leqslant (1+\epsilon)|E|_e - \frac{1}{2\cdot 5^n}|E|_e = \left(1 + \epsilon - \frac{1}{2\cdot 5^n}\right)|E|_e \leqslant \left(1 - \frac{1}{5^{n+1}}\right)|E|_e.$$

Now, cover this portion using the Simple Vitali Lemma again. Doing so, we get a sequence $\{Q_i\}_{i=N_1+1}^{N_2}$, and we have

$$\left|E - \bigcup_{i=1}^{N_2} Q_i\right| \leqslant \left(1 - \frac{1}{5^{n+1}}\right)\left|E - \bigcup_{i=1}^{N_1} Q_j\right| \leqslant \left(1 - \frac{1}{5^{n+1}}\right)^2 |E|_e.$$

Since $\left(1-\left(1/5^{n+1}\right)\right)^k \to 0$ as $k \to \infty$, we find that

$$\left|E - \bigcup_{j=1}^{\infty} Q_j\right| = 0,$$

thus giving us the first property. **Q.E.D**

We can then use this to get a nice Corollary on finite collections.

Corollary 5.5.1. Under the same assumptions as the Vitali Covering Lemma **Theorem 5.5**, there exists a finite sequence $Q_1, Q_2, \ldots, Q_N$ of disjoint cubes such that

(i)

$$\left|E - \bigcup_{j=1}^{N} Q_j\right| < \epsilon$$

(ii)

$$\sum_j |Q_j| < (1+\epsilon)|E|_e$$

(iii)

$$\sum_{j=1}^{N} |Q_j| \geqslant \left|E \cap \left(\bigcup_{j=1}^{N} Q_j\right)\right| > (1-\epsilon)|E|_e.$$

Remark 27. The third property is the one of interest, as the other two are really just results from **Theorem 5.5**. It turns out that it is also a direct result of (i) and (ii).

Proof. First, note that

$$\sum_{j=1}^{N} |Q_j| \geqslant \left|E \cap \left(\bigcup_{j=1}^{N} Q_j\right)\right|.$$

Next, we note that

$$|E|_e = \left|E \cap \left(\bigcup_{j=1}^{N} Q_j\right)\right|_e + \left|E - \bigcup_{j=1}^{N} Q_j\right|_e$$

by **Theorem 2.4**. Thus, using (i), we get the desired result. **Q.E.D**

5.5 Lecture 22 (Differentiability of Monotone Functions)

Theorem 5.6. If $f(x)$ is non-decreasing (the book uses increasing) on (a, b), then f has a measurable derivative $f'(x)$ on $[a, b]$, and furthermore

$$0 \leqslant \int_a^b f'(x)dx \leqslant f(b-) - f(a+),$$

where $f(b-) = \lim_{x \to b^-} f(x)$ and $f(a+) = \lim_{x \to a^+} f(x)$.

Remark 28. The fundamental theorem of Calculus gives us the idea that we should expect equality. It is somewhat surprising that we, in fact, do **not** have equality always. We will explore why later this lecture.

Proof. There are four "derivatives" we need to check:

$$D_1 f(x) = \limsup_{h \to 0^+} \frac{f(x+h) - f(x)}{h},$$

$$D_2 f(x) = \liminf_{h \to 0^+} \frac{f(x+h) - f(x)}{h},$$

$$D_3 f(x) = \limsup_{h \to 0^-} \frac{f(x+h) - f(x)}{h},$$

$$D_4 f(x) = \liminf_{h \to 0^-} \frac{f(x+h) - f(x)}{h}.$$

So we need to have $D_1 = D_2 = D_3 = D_4$ a.e. on $[a, b]$. Let's start with $D_1 f = D_4 f$ a.e. In order to do this, we'll show

$$|\{D_1 f > D_4 f\}| = 0$$

and

$$|\{D_4 f > D_1 f\}| = 0.$$

First, let's note that

$$\{D_1 f > D_4 f\} = \bigcup_{r,s \in \mathbb{Q}} \{D_1 f > r > s > D_4 f\}.$$

So we just need to show that for all $r, s \in \mathbb{Q}$,

$$|\{D_1 f > r > s > D_4 f\}| = 0.$$

Fix r, s, and let $A = \{D_1 f > r > s > D_4 f\}$. Assume for contradiction that $|A| > 0$. Since $D_4(x) < s$ on A, we can cover the set A by intervals $[x - h, x]$ such that

$$\frac{f(x) - f(x-h)}{h} < s$$

for h arbitrarily small. We rewrite it as $f(x) - f(x-h) < sh$. So these cubes that cover A cover it in a Vitali sense. Using the Vitali Covering Lemma (**Theorem 5.5**), we get that there exist intervals $[x_i - h_i, x_i]$ for $i = 1, \ldots, N$ with

(i) $f(x_i) - f(x_i - h_i) < sh_i$,

(ii) $\sum_i h_i \leqslant (1+\epsilon)|A|$ for $\epsilon > 0$ fixed.

So this gives us some information on the left derivative. Let's now focus on the right derivative. Let

$$B = A \cap \left(\bigcup_{i=1}^{N} [x_i - h_i, x_i) \right).$$

The set B is going to have finite positive measure. That is, we have $|B| > |A| - \epsilon > 0$ (see **Corollary 5.5.1**). The Vitalli Covering lemma implies there exist $[y_j, y_j + h_j]$, $j = 1, \ldots, M$, such that

(i) $f(y_j + h_j) - f(y_j) > rh_j$,

(ii) $\sum_j h_j > |B| - \epsilon$,

(iii) $[y_j, y_j + h_j] \leqslant [x_i - h_i, x_i]$ for some i.

Stringing these facts together, we have

$$\sum_{j=1}^{M} f(y_j + h_j) - f(y_j) > r \sum_{j=1}^{M} h_j > r(|B| - \epsilon) \geqslant r(|A| - 2\epsilon).$$

Now, we use monotonicity to note that

$$\sum_{j=1}^{M} f(y_j + h_j) \leqslant \sum_{i=1}^{N} f(x_i) - f(x_i - h_i) < s \sum_{i=1}^{N} h_i < s(1+\epsilon)|A|.$$

Putting things together, we have

$$r(|A| - 2\epsilon) < s(1+\epsilon)|A|.$$

However, we assumed that $s < r$. So this results in a contradiction for ϵ sufficiently small.

Now, we just need to show that

$$0 \leqslant \int_a^b f'(x)dx \leqslant f(b-) - f(a+).$$

We're going to extend the definition of $f(x)$ so that $f = f(b-)$ for $x \geqslant b$. We do this to form

$$f_k(x) = \frac{f(x + 1/k) - f(x)}{1/k},$$

where $f_k \to f'$ a.e. in (a, b). Furthermore,

$$0 \leqslant f_k(x) \leqslant f'$$

a.e. in (a,b). So we get

$$0 \leqslant \int_a^b f'(x)dx.$$

Now we use Fatou's lemma (**Theorem 4.9**) to see

$$\int_a^b f'(x)dx \leqslant \liminf_{k\to\infty} \int_a^b f_k(x)dx.$$

Notice that on the right hand side we have

$$\liminf_{k\to\infty} \int_a^b f_k(x)dx = \liminf_{k\to\infty} \int_a^b \frac{f(x+1/k)-f(x)}{1/k}dx = \liminf_{k\to\infty} \int_a^b \frac{f(x+1/k)}{1/k}dx - \liminf_{k\to\infty} \int_a^b \frac{f(x)}{1/k}$$

$$= \liminf_{k\to\infty} k\left(\int_{a+1/k}^{b+1/k} f(x)dx - \int_a^b f(x)dx\right) = \liminf_{k\to\infty} k\left(\int_b^{b+1/k} f(x)dx - \int_a^{a+1/k} f(x)dx\right)$$

$$\liminf_{k\to\infty} k\left(\frac{f(b-)}{k} - \frac{f(a+)}{k}\right) = f(b-) - f(a+),$$

since f is monotone, $f(x) \approx f(a+)$ in $(a, a+1/k)$.

Remark 29. I think you can simplify this using the fact that $f(a+) \leqslant f(a)$ for $x \in (a,b)$ to just get a bound, but he did it this way.

Q.E.D

We now explore some examples where there is not equality.

Example 5.2. Let $f(x)$ be the Cantor function. Then

$$\int_0^1 f'(x)dx = 0,$$

since $f'(x) = 0$ a.e. However, $f(1) - f(0) = 1$. So we do not have equality, and the fundamental theorem of Calculus fails.

We now want to determine when exactly the FTC will hold. We will first need a definition.

Definition. A function f is **absolutely continuous** on $[a,b]$ if, given $\epsilon > 0$, there exists a $\delta > 0$ such that if $\{[a_i,b_i]\}_i$ is a countable family of non-overlapping subintervals of $[a,b]$, then

$$\sum_i |f(b_i) - f(a_i)| < \epsilon$$

if

$$\sum_i |b_i - a_i| < \delta.$$

It turns out that this is a sufficient condition for the FTC.

Theorem 5.7. We have that f is absolutely continuous on $[a,b]$ if and only if f' exists a.e. and $f(x) = f(a) + \int_a^x f'(y)dy$ for $x \in [a,b]$.

5.6 Lecture 23 (Conditions for FTC to hold pt. 1)

Remark 30. It turns out that this definition of absolutely continuous aligns with our definition of absolutely continuous for set functions.

Below we explore some examples of absolutely continuous functions.

Example 5.3. (i) The **indefinite integral**

$$F(A) = \int_A g(x)dx$$

is an absolutely continuous function. Let's say

$$f(x) = \int_a^x g(t)dt.$$

Then we can see that $f(x)$ is absolutely continuous on some interval $[a, b]$, since

$$f(b_i) - f(a_i) = \int_{[a_i,b_i]} g(t)dt = F([a_i, b_i]).$$

(ii) A **Lipschitz function** is clearly absolutely continuous.

(iii) A non-example is the **Cantor function**.

Lemma 5.4. If f is absolutely continuous, then $f \in BV([a, b])$, where $BV([a, b])$ is the set of all functions with bounded variation on $[a, b]$.

Remark 31. This will imply the derivative exists almost everywhere. We have that $f \in BV([a, b])$ implies that $f = P - N$, where P and N are monotone non-decreasing functions, and last week tells us the derivatives exists a.e. for monotone functions. So $f' = P' - N'$ exists a.e.

Proof. Since we have absolutely continuous, let's fix a $\delta > 0$ so that if

$$\sum_i (b_i - a_i) < \delta$$

are non-overlapping, then

$$\sum_i |f(b_i) - f(a_i)| < 1.$$

Now, divide $[a, b]$ into N intervals of length $\frac{b-a}{N}$, where we make N large enough so that $\frac{b-a}{N} < \delta$. Then we have that

$$Var(f; [a, b]) = \sum_{i=1}^{N} Var\left(f; \left[a + \frac{i-1}{N}(b-a), a + \frac{i}{N}(b-a)\right]\right)$$

and we have that

$$Var\left(f;\left[a+\frac{i-1}{N}(b-a),a+\frac{i}{N}(b-a)\right]\right)\leqslant 1.$$

Therefore, we get

$$Var(f;[a,b])\leqslant N<\infty.$$

Q.E.D

Definition. A function f is **singular** on $[a,b]$ if $f'=0$ a.e. on $[a,b]$.

Remark 32. From Calculus, one might expect singular functions to always be constant, but the Cantor function is singular.

Lemma 5.5. If f is absolutely continuous and singular on $[a,b]$, then f is constant.

Proof. It is enough to show that $f(a)=f(b)$, since we can apply the same argument for $[a,x]$, $x\in(a,b]$. Fix some $\epsilon>0$, then there exists a $\delta>0$ such that

$$\sum_i(b_i-a_i)<\delta \implies \sum_i|f(b_i)-f(a_i)|<\epsilon.$$

Examine $E=\{x\ :\ f'(x)=0\}$. Note that $|E|=b-a$. Thus, for every $x\in E$, we can find an $h>0$ such that $|f(x+h)-f(x)|<\epsilon h$. So we apply the Vitali Covering Lemma (**Theorem 5.5**) to get $\{[x_i,x_i+h_i]\}_{i=1}^N$ such that

(i)

$$|f(x_i+h_i)-f(x_i)|<\epsilon h_i,$$

(ii)

$$\sum_i h_i>(b-a)\delta.$$

Remark 33. Note here that $\sum_{i=1}^N|f(x_i+h_i)-f(x_i)|\leqslant\epsilon\sum_{i=1}^N h_i\leqslant\epsilon(b-a)$.

The complement of $[x_i,x_i+h_i]$ is non-overlapping intervals of total length less than δ. So the increments on this are less than ϵ. So this means that $|f(b)-f(a)|\leqslant$ (increments on $[x_i,x_i+h_i]$)+ (increments not in $[x_i,x_i+h_i]$) $\leqslant\epsilon(b-a)+\epsilon$. Let $\epsilon\to0$, and we win. **Q.E.D**

5.7 Lecture 24 (Conditions for FTC to hold pt. 2, Convex functions)

We now finally prove **Theorem 5.7**.

Proof. ($\Longleftarrow$) We established in the prior lecture that $x \mapsto \int_a^x f'(y)dy$ is absolutely continuous (see **Example 5.3**). We therefore get that F is absolutely continuous.
($\Longrightarrow$) We showed last time that f' exists exists a.e. (see **Lemma 5.4** and **Remark 31**). Let's suppose $F(x) = \int_a^x f'(y)dy$. We know that $F(x)$ is absolutely continuous and $F' = f'$ a.e. by Lebesgue differentiation. If we look at $F - f$, we see that $(F - f)' = F' - f' = 0$, so that this is singular. Hence, $F - f$ is constant. Now, examine $F(x) - f(x) = F(a) - f(a)$. We have that $F(a) = 0$, and we have

$$F(x) - f(x) = \int_a^x f'(y)dy - f(x) = -f(a).$$

Rearranging terms gives us

$$\int_a^x f'9y) = f(x) - f(a).$$

Q.E.D

Thus, we now have the condition where the FTC holds!

Corollary 5.7.1. If $f \in BV([a, b])$, then $f = g + h$, where g is absolutely continuous and h is singular. Furthermore, this representation is unique up to constants.

Proof. $f \in BV([a, b])$ implies that f' exists a.e. and $f' \in L([a, b])$. Now, we can set up the function

$$g(x) = \int_a^x f'(y)dy,$$

which is absolutely continuous. Let

$$h(x) = f(x) - g(x).$$

Then we have

$$h'(x) = f'(x) - g'(x) = f'(x) - f'(x) = 0.$$

So $h(x)$ must be singular. To see that this is unique up to constants, use the prior theorem. **Q.E.D**

Chapter 6

Inequalities and L^p spaces

We first want to discuss convex functions.

Definition. A function ϕ is **convex** on (a, b) if

$$\phi(tx_0 + (1-t)x_1) \leqslant t\phi(x_0) + (1-t)\phi(x_1)$$

for all $x_0, x_1 \in (a, b)$ and for all $t \in [0, 1]$.

Geometrically, there's a nice way of viewing convex functions. We have that the line segment connecting $\phi(x_0)$ to $\phi(x_1 0$ always lies above the graph.

Theorem 6.1. If ϕ' exists and is non-decreasing on (a, b), then ϕ is convex. In particular, if $\phi'' \geqslant 0$, then ϕ is convex.

Using this theorem, we can create some examples of convex functions.

Example 6.1. (i) $\phi(x) = e^{ax}$ is convex.

(ii) $\phi(x) = x^p$ is convex for $p \geqslant 1$.

(iii) $\phi(x) = |x|^p$ is convex.

(iv) $\phi(x) = -\log(x)$ for $x > 0$ is convex.

Proof. Fix $x_0 < x_1$ in (a, b). Let $x_t = (1-t)x_0 + tx_1$. We do so in order to have $x_t|_{t=0} = x_0$ and $x_t|_{t=1} = x_1$. In order to prove this, we need to show that $\phi(x_t) \leqslant t\phi(x_1) + (1-t)\phi(x_0)$. On the boundary, this condition is easy to show, and so we restrict ourselves to viewing $x_0 < x_t < x_1$. We then use the mean value theorem, since ϕ' exists. That is, we have ξ_1, ξ_1 such that $x_0 < \xi_1 < x_t < \xi_2 < x_1$, and

$$\phi'(\xi_1) = \frac{\phi(x_t) - \phi(x_0)}{x_t - x_0},$$

$$\phi'(\xi_2) = \frac{\phi(x_1) - \phi(x_t)}{x_1 - x_t}.$$

We can simplify this to

$$\phi'(\xi_1) = \frac{\phi(x_t) - \phi(x_0)}{t(x_1 - x_0)},$$

$$\phi'(\xi_2) = \frac{\phi(x_1) - \phi(x_t)}{(1-t)(x_1 - x_0)}.$$

Since ϕ' is non-decreasing, we get

$$\phi'(\xi_1) \leqslant \phi'(\xi_2),$$

which tells us

$$\frac{\phi(x_t) - \phi(x_0)}{t(x_1 - x_0)} \leqslant \frac{\phi(x_1) - \phi(x_t)}{(1-t)(x_1 - x_0)}.$$

Simplifying, we just get

$$\frac{\phi(x_t) - \phi(x_0)}{t} \leqslant \frac{\phi(x_1) - \phi(x_t)}{(1-t)},$$

which reduces down to

$$\phi(x_t) \leqslant t\phi(x_1) + (1-t)\phi(x_0),$$

as desired. **Q.E.D**

Now we want to explore what you can say about a function if it is convex.

Theorem 6.2. If ϕ is convex on (a, b), then

(i) ϕ is continuous on (a, b),

(ii) $\phi'(x)$ exists at all but at most countably many points in (a, b), and is non-decreasing.

Proof. Let

$$D^+\phi(x) = \lim_{h \to 0^+} \frac{\phi(x+h) - \phi(x)}{h},$$

$$D^-\phi(x) = \lim_{h \to 0^+} \frac{\phi(x) - \phi(x+h)}{h}.$$

Convexity tells us that $D^+\phi(x)$ is decreasing in h, and $D^-\phi(x)$ is increasing in h. Since these are monotone and bounded, we have that the limits exist everywhere. Using convexity, we can also see that

$$-\infty < D^-\phi(x) \leqslant D^+\phi(x) < \infty.$$

This is then enough to prove that ϕ is continuous, since if it wasn't continuous we'd have a discontinuity with a non-finite derivative, which is a contradiction. For any $x < y$, we have

$$D^+\phi(x) \leqslant \frac{\phi(y) - \phi(x)}{y - x} \leqslant D^-\phi(y).$$

Therefore, we get that

$$D^-\phi(x) \leqslant D^+\phi(x) \leqslant D^-\phi(y) \leqslant D^+\phi(y).$$

Note that $D^+\phi$ and $D^-\phi$ are monotone, and so we have that (using a property about monotone functions) there are only countably many discontinuities. Suppose x is a point where they are both continuous, Then we get equality by dragging the limit in; that is

$$D^+\phi(x) \leqslant \lim_{y\to x} D^-\phi(y) = D^-\phi(x).$$

So the derivative exists at all but countably many points. **Q.E.D**

Corollary 6.2.1. If ϕ convex on (a,b), then ϕ is Lipschitz on $[x_1, x_2] \subseteq (a,b)$, and, in particular,

$$\phi(x_2) - \phi(x_1) = \int_{x_1}^{x_2} \phi'(x)dx.$$

6.1 Lecture 25 (Inequalities (Jensen, Holder, Young))

There are two big inequalities we'd like to discuss: Jensen's inequality (both discrete and continuous) and Holder's inequality. We'll first talk about Jensen's inequality, and specifically the discrete case.

Theorem 6.3. (Jensen's Inequality (Discrete)) Let ϕ be a convex function on (a,b). If we have $x_1,\ldots,x_n \in (a,b)$ and $t_1,\ldots,t_n \geqslant 0$ such that $\sum_{i=1}^n t_i = 1$, then

$$\phi\left(\sum_{i=1}^n t_i x_i\right) \leqslant \sum_{i=1}^n t_i\phi(x_i).$$

Proof. The proof is done via induction. The cases $n = 1$ and 2 are not interesting, and so we'll jump to the case $n = 3$. With $n = 3$, we want to examine $\phi(x_1t_1 + x_2t_2 + x_3t_3)$. Rewrite this as

$$\phi\left(t_1x_1 + (1-t_1)\frac{t_2x_2 + t_3x_3}{1-t_1}\right).$$

Using the definition of convexity, we get

$$\phi\left(t_1x_1 + (1-t_1)\frac{t_2x_2 + t_3x_3}{1-t_1}\right) \leqslant t_1\phi(x_1) + (1-t_1)\phi\left(\frac{t_2x_2 + t_3x_3}{1-t_1}\right).$$

Use it again, noting that $\frac{t_2+t_3}{1-t_1} = 1$ to get that

$$t_1\phi(x_1)+(1-t_1)\phi\left(\frac{t_2x_2 + t_3x_3}{1-t_1}\right) \leqslant t_1\phi(x_1)+(1-t_1)\left(\frac{t_2}{1-t_1}\phi(x_2) + \frac{t_3}{1-t_1}\phi(x_3)\right)$$
$$= t_1\phi(x_1) + t_2\phi(x_2) + t_3\phi(x_3).$$

For induction, we just abuse this type of trick. **Q.E.D**

We now look at the "continuous" case.

Theorem 6.4. (Jensen's Inequality (Continuous)) If $p(x) \geqslant 0$ and $\int_A p(x)dx = 1$ and ϕ is convex on $f(A) \subseteq (a,b)$, then

$$\phi\left(\int_A f(x)p(x)dx.\right) \leqslant \int_A \phi(f(x))p(x)dx.$$

Before proving this, we want to explore a special case which may make this easier to remember.

Example 6.2. Let $p(x) = \frac{1}{|A|} \cdot \chi_A$. Then Jensen's Inequality says

$$\phi\left(\frac{1}{|A|}\int_A f(x)dx\right) \leqslant \frac{1}{|A|}\int_A \phi(f(x))dx.$$

That is, ϕ(Average of f) $\leqslant$ Average of $\phi(f)$. Therefore, we have

$$(\text{Average of f})^2 \leqslant \text{Average of } f^2.$$

From probability, we know that

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \geqslant 0,$$

which tells us the same thing.

Proof. Let $\gamma = \int_A f(x)p(x)dx$ for notational simplicity. Then we have $\alpha \leqslant \gamma \leqslant b$. Since ϕ is convex, there exists a "supporting line," which is almost what we would define to be a tangent line. The issue is that, as discussed in **Theorem 5.9 (ii)**, there are countably many points where $\phi'(x)$ does not exist. At these points, there is no such thing as well defined supporting line, so to remediate that we just define a supporting line to be a line that is through the point and is always less than or equal to ϕ, and we'll just take one of these possibilities. Therefore, we have that the supporting line looks something like $y = m(x-\gamma) + \phi(\gamma)$., and the fact that it is a supporting line tells us that

$$m(x-\gamma) + \phi(\gamma) \leqslant \phi(x)$$

for all $x \in (a,b)$. Now, we can say that

$$m(f(x)-\gamma) + \phi(\gamma) \leqslant \phi(f(x)),$$

since $f(x) \in (a,b)$. Taking integrals of both sides and multiplying by $p(x)$ preserves this inequality, and so we get

$$\int_A mp(x)f(x)dx - \int_A mp(x)\gamma dx + \int_A p(x)\phi(\gamma) \leqslant \int_A \phi(f(x))p(x)dx.$$

Recall that we forced

$$\int_A p(x) = 1,$$

and so moving things around we have

$$m\int_A p(x)f(x)dx - m\gamma + \phi(\gamma) \leqslant \int_A \phi(f(x))p(x)dx.$$

By how we defined γ, we can rewrite this as

$$m\gamma - m\gamma + \phi(\gamma) = \phi(\gamma) = \phi\left(\int_A p(x)f(x)dx\right) \leqslant \int_A \phi(f(x))p(x)dx.$$

Q.E.D

We mention briefly another inequality.

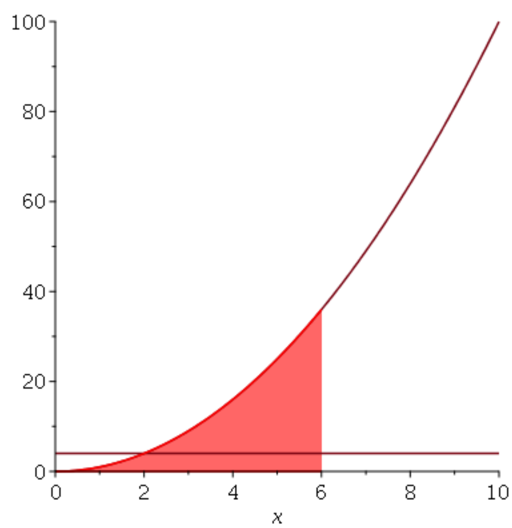
Lemma 6.1. (Young's Inequality) If $a, b \geqslant 0$, $p, q \geqslant 1$, $1/p + 1/q = 1$, then

$$ab \leqslant \frac{1}{p}a^p + \frac{1}{q}b^q.$$

Proof. This will be done via areas of a graph. Let $y = x^{p-1}$ be a graph, then we see that $x = y^{1/(p-1)}$. Pick arbitrary a on the x-axis and arbitrary b on the y-axis. Then we have that the area from 0 to a of y is

$$\int_0^a x^{p-1}dx = \frac{1}{p}a^p.$$

That is, the area in the graph below is $1/pa^p$;



Likewise, finding the area from 0 to b of x, we have

$$\int_0^b y^{\frac{1}{p-1}}dy = \frac{p-1}{p}b^{\frac{p}{p-1}}.$$

Notice that since

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

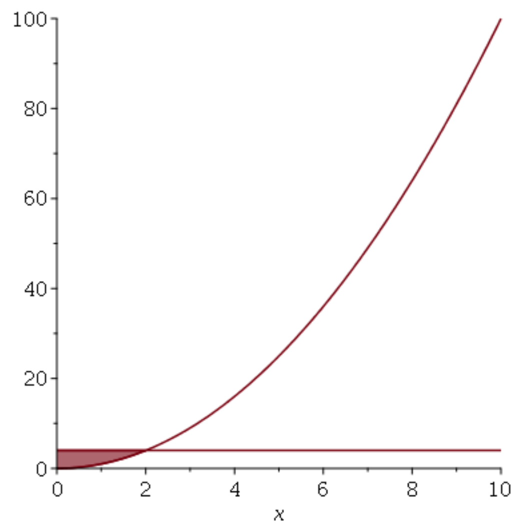
we have

$$\frac{1}{q} = 1 - \frac{1}{p} = \frac{p-1}{p},$$

so that the area is

$$\frac{1}{q}b^q.$$

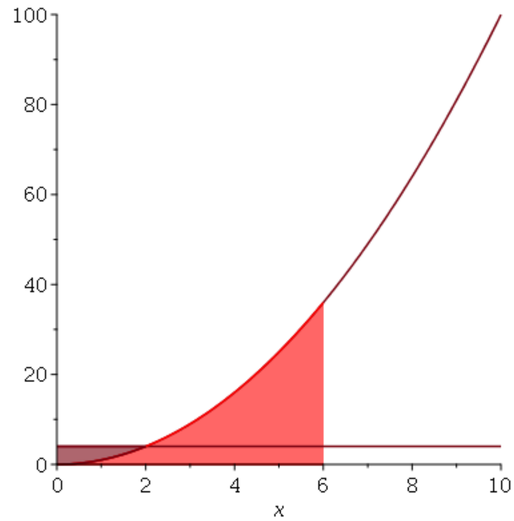
This is the area in the graph below;



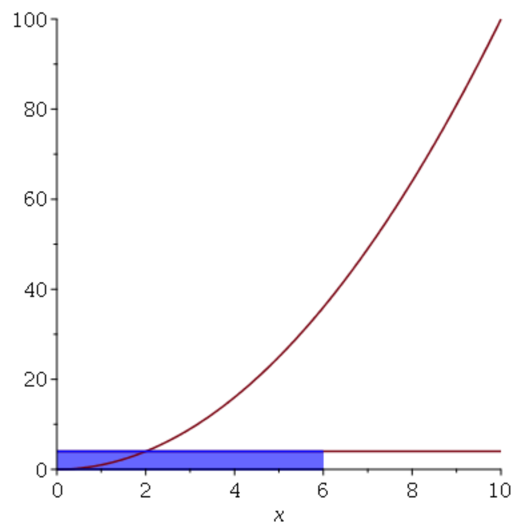
So putting this together, we have that the total area is going to be

$$\frac{1}{p}a^p + \frac{1}{q}b^q,$$

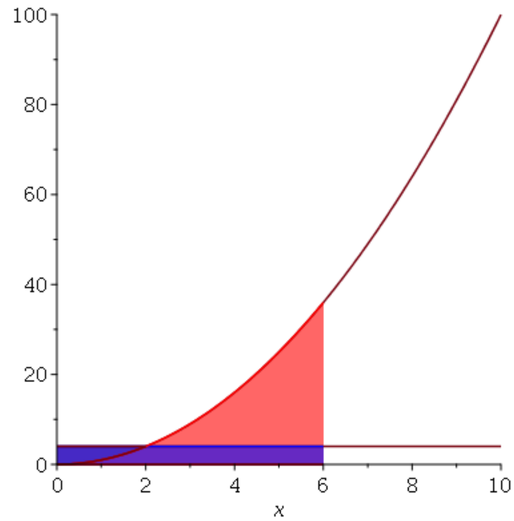
which corresponds to the area of the graph below;



Now, compare this to the area of the box with height b and width a. This would simply be ab, which is the area of the graph given below;



Comparing this to the graph prior, we see that it is contained in it. Visually, we have



But notice that this means that

$$ab \leqslant \frac{1}{p}a^p + \frac{1}{q}b^q,$$

as desired. This does not rely on picking a b smaller than our a either; we can follow the same procedure and get the same result. **Q.E.D**

We now use this inequality to prove Holder's Inequality.

Theorem 6.5. (Holder's Inequality) If $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\int |fg| \leqslant \left(\int |f|^p \right)^{1/p} \left(\int |g|^q \right)^{1/q}.$$

Remark 34. We will eventually show that we can let $p = 1$ and $p = \infty$ and get the same result.

Proof. First, let's suppose that

$$\int |f|^p = 1, \quad \int |g|^q = 1.$$

Note that

$$|fg| \leqslant \frac{1}{p}|f|^p + \frac{1}{q}|g|^q$$

by Young's Inequality. Then we can integrate both sides to get

$$\int |fg| \leqslant \int \left(\frac{1}{p}|f|^p + \frac{1}{q}|g|^q \right).$$

Using linearity and pulling out constants, we have

$$\int |fg| \leqslant \frac{1}{p}\int |f|^p + \frac{1}{q}\int |g|^q = \frac{1}{p} + \frac{1}{q} = 1 = \left(\int |f|^p\right)^{1/p}\left(\int |g|^q\right)^{1/q}.$$

Now, it may seem that this was a restrictive case and so pointless, however we will see that this captures all functions. Now, take f and g to be general functions. Consider

$$\widetilde{f}(x) = \frac{f(x)}{\left(\int |f|^p\right)^{1/p}}.$$

Since

$$\left(\int |f|^p\right)^{1/p}$$

is just going to be some constant, let's rewrite

$$\widetilde{f}(x) = Af(x).$$

Define $\widetilde{g}(x)$ analogously. Notice that we have

$$\int |\widetilde{f}|^p = A^p \int |f|^p = \frac{\int |f|^p}{\int |f|^p} = 1,$$

and likewise

$$\int |\widetilde{g}|^q = B^q \int |g|^q = 1.$$

So for $\widetilde{f}$, $\widetilde{g}$, we have that Holder's inequality holds by our prior work. Now, let's notice that

$$\int |\widetilde{f}\widetilde{g}| = \int AB|fg| = AB\int |fg|.$$

Likewise, notice that

$$\left(\int |\widetilde{f}|^p\right)^{1/p}\left(\int |\widetilde{g}|^q\right)^{1/q} = AB\left(\int |f|^p\right)^{1/p}\left(\int |g|^q\right)^{1/q}.$$

We thus find that

$$\int |\widetilde{f}\widetilde{g}| \leqslant \left(\int |\widetilde{f}|^p\right)^{1/p}\left(\int |\widetilde{g}|^q\right)^{1/q} \leftrightarrow \int |fg| \leqslant \left(\int |f|^p\right)^{1/p}\left(\int |g|^q\right)^{1/q}.$$

Q.E.D

This leads us to a very famous Corollary.

Corollary 6.5.1. (Cauchy-Schwartz Inequality) We have

$$\int |fg| \leqslant \sqrt{\int f^2 \int g^2}.$$

We now want to reword this in terms of L^p norms, but to do that we need to first understand an L^p space.

Definition. We have that the **L^p space over E** is

$$L^p(E) = \left\{ f \ : \ \int_E |f|^p < \infty \right\}.$$

Definition. We define the **L^p norm** to be

$$||f||_p = ||f||_{p,E} = \left(\int_E |f|^p \right)^{1/p}.$$

We generally drop the E since it's understood in context.

One thing we've seen/used is that

$$||cf||_p = c||f||_p.$$

We may now also reword the L^p space definition. That is, the L^p space over E is

$$L^p(E) = \{ f \ : \ ||f||_p < \infty \}.$$

We can also define the $L^\infty(E)$ space to be the set of bounded functions. The $L^\infty(E)$ norm, $||f||_\infty$, is the essential supremum of f over E. That is,

$$||f||_\infty = \inf\{M \ : \ |f(x)| \leqslant M \text{ a.e. on E}\} = \text{the essential supremum of } f.$$

Going back to Holder's inequality, we may rewrite it in the form

$$||fg||_1 \leqslant ||f||_p ||g||_q.$$

With this, it makes sense to also define things when $p = \infty$ and $p = 1$. Assume wlog that $p = \infty$, then $q = 1$. Therefore, we have

$$\int |fg| \leqslant \int ||f||_\infty |g| \leqslant ||f||_\infty \int |g| = ||f||_\infty ||g||_1.$$

6.2 Lecture 26 (L^p space structure)

One may ask why we define things differently when $p = \infty$. The reason is that this adheres to limits.

Theorem 6.6. If $|E| < \infty$, then $||f||_\infty = \lim_{p\to\infty} ||f||_p$.

Proof. Fix $a < ||f||_\infty$ and consider $E_a = \{|f| > a\}$. We can use Chebychev's Inequality (**Theorem 4.4**) to note that

$$|E_a| \leqslant \frac{1}{a} \int_E f.$$

Equivalently, we may write $E_a = \{|f|^p > a^p\}$. Then we have

$$|E_a| \leqslant \frac{1}{a^p}\int_E f^p \leftrightarrow a^p|E_a| \leqslant \int_E f^p.$$

Now, take both sides to the $1/p$ power to get

$$(a^p|E_a|)^{1/p} = a|E_a|^{1/p} \leqslant \left(\int_E f^p\right)^{1/p} = ||f||_p.$$

We note that $|E_a| < \infty$, and we will need a result from Freshman Calculus.

Claim 6.1. If $c \in \mathbb{R}_{>0}$, then

$$\lim_{x\to\infty} c^{1/x} = 1.$$

Proof. Let $y = c^{1/x}$. Then taking the log of both sides, we have

$$\log(y) = \frac{\log(c)}{x}.$$

Taking the limit as $x \to \infty$ gives

$$\lim_{x\to\infty} \log(y) = 0.$$

Hence, we have

$$\lim_{x\to\infty} e^{\log(y)} = \lim_{x\to\infty} c^{1/x} = e^0 = 1.$$

Q.E.D

Since $a < ||f||_\infty$, we get that $|E_a| > 0$ (provided $|E| \neq 0$ and $||f||_\infty \neq 0$). Hence, we have

$$a \leqslant \liminf_{p\to\infty} ||f||_p.$$

Now, since this works for all $a < ||f||_\infty$, we get that

$$||f||_\infty \leqslant \liminf_{p\to\infty} ||f||_p.$$

Next, we notice that for all p we have

$$||f||_p \leqslant ||f||_\infty |E|^{1/p}.$$

So, as $p \to \infty$, we get

$$\limsup_{p\to\infty} ||f||_p \leqslant ||f||_\infty.$$

Chaining things together, we get

$$||f||_\infty \leqslant \liminf_{p\to\infty} ||f||_p \leqslant \limsup_{p\to\infty} ||f||_p \leqslant ||f||_\infty.$$

Hence,

$$\lim_{p\to\infty} ||f||_p = ||f||_\infty.$$

Q.E.D

Theorem 6.7. If $|E| < \infty$, $0 < p_1 < p_2 \leqslant \infty$, then

$$L^{p_2}(E) \subseteq L^{p_1}(E).$$

Proof. Notice that

$$\int_E |f|^{p_1} = \int_E |f|^{p_1} \cdot \chi_E.$$

Now, we use Hölder's inequality with $p = p_2/p_1$ and $1/q = (p_2 - p_1)/p_2$ to get

$$\int_E |f|^{p_1} \cdot \chi_E \leqslant \left(\int_E |f|^{p_2}\right)^{p_1/p_2} \left(\int_E |\chi_E|^q\right)^{1/q} = ||f||_{p_2}^{p_1} |E|^{1/q}.$$

Since $|E| < \infty$, we have $|E|^{1/q} < \infty$. If $||f||_{p_2} < \infty$, then we get for free that $||f||_{p_2}^{p_1} < \infty$, and the above inequality tells us that $||f||_{p_1} < \infty$. **Q.E.D**

We want to now talk about the structure of these L^p spaces.

Definition. A **Banach space** is a vector space (over either $\mathbb{R}$ or $\mathbb{C}$) with a norm that is complete. In shorter words, it is a complete normed vector space.

See the **Chapter 1** for the definition of **vector space**, **complete space**, and **norm space**.

Remark 35. Banach spaces are the prototypical example of infinite dimensional vector spaces.

Theorem 6.8. $L^p(E)$ is a Banach space for any $1 \leqslant p \leqslant \infty$ over $\mathbb{R}$ or $\mathbb{C}$ on the equivalence classes of $\sim$, where $f \sim g$ if and only if $f = g$ a.e.

Proof. We break this up into each aspect. Throughout, let $F = \mathbb{R}$ or $\mathbb{C}$ be the field which this space is over.
Vector space: If $f \in L^p$, then for $c \in F$ we have $cf \in L^p$. This is because $f \in L^p$ implies

$$\left(\int_E |c^p f^p|\right)^{1/p} = |c| \left(\int_E |f|^p\right)^{1/p} < \infty.$$

For the closure under addition, we will need a result.

Lemma 6.2. For any $p \geqslant 1$ and $n \geqslant 1$, there exists a constant $C_{n,p} > 0$ such that

$$\left(\sum_{k=1}^n a_k\right)^p \leqslant C_{n,p} \sum_{k=1}^n a_k^p,$$

for any non-negative numbers $a_1, \ldots, a_n$.

Proof. **Example 5.4 (ii)** says that $\phi(x) = x^p$ is a convex function if $p \geqslant 1$. **Theorem 5.10** then tells us

$$\phi\left(\frac{\sum_{k=1}^n a_k}{n}\right) \leqslant \frac{\sum_{k=1}^n \phi(a_k)}{n}.$$

Hence, we have

$$\frac{\left(\sum_{k=1}^{n} a_k\right)^p}{n^p} \leqslant \frac{\sum_{k=1}^{n} a_k^p}{n}.$$

So, rewriting this, we get

$$\left(\sum_{k=1}^{n} a_k\right)^p \leqslant n^{p-1} \sum_{k=1}^{n} a_k^p.$$

Q.E.D

Now, using **Lemma 5.7**, we get that

$$|f+g|^p \leqslant 2^{p-1}\left(|f|^p + |g|^p\right).$$

This tells us that $f, g \in L^p$, then $f+g \in L^p$ for $1 \leqslant p < \infty$ by integrating and using the linearity of integration. For $p = \infty$, we clearly get

$$||f+g||_\infty \leqslant ||f||_\infty + ||g||_\infty.$$

Thus, it is a vector space.
Norm: Notice from above that

$$||cf||_p = |c|||f||_p.$$

Now, we need to establish $||f||_p = 0$ if and only if $f = 0$ a.e. This follows from **Theorem 4.2 (viii)**. Finally, we need to establish the triangle inequality; that is,

$$||f+g||_p \leqslant ||f||_p + ||g||_p.$$

It turns out that this is an important inequality known as Minkowski's inequality.

Theorem 6.9. (Minkowski Inequality) If $1 \leqslant p \leqslant \infty$,

$$||f+g||_p \leqslant ||f||_p + ||g||_p.$$

Proof. **Case** 1: Let $p = 1$. Then we have the normal triangle inequality;

$$|f+g| \leqslant |f| + |g|.$$

Integrate both sides to get

$$||f+g||_1 \leqslant ||f||_1 + ||g||_1.$$

Case 2: Let $p = \infty$. Then we have

$$|f(x)+g(x)| \leqslant |f(x)| + |g(x)| \leqslant ||f||_\infty + ||g||_\infty \text{ a.e.}$$

Therefore, the least upper bound property gives us

$$||f+g||_\infty \leqslant ||f||_\infty + ||g||_\infty.$$

Case 3: Now, let $1 < p < \infty$. Then we have

$$||f+g||_p^p = \int_E |f+g|^p = \int_E |f+g|^{p-1}|f+g|.$$

The triangle inequality gives us

$$\int_E |f+g|^{p-1}|f+g| \leqslant \int_E |f+g|^{p-1}|f| + \int_E |f+g|^{p-1}|g|.$$

Now, Hölder's inequality with $p' = p/(p-1)$ and $q = p$ gives us

$$\int_E |f+g|^{p-1}|f| \leqslant \left(\int_E |f+g|^p\right)^{(p-1)/p} \left(\int_E |f|^p\right)^{1/p} + \left(\int_E |f+g|^p\right)^{(p-1)/p} \left(\int_E |g|^p\right)^{1/p}$$

$$\leqslant ||f+g||_p^{p-1}\left(||f||_p + ||g||_p\right).$$

Dividing both sides by $||f+g||_p^{p-1}$ gives

$$||f+g||_p \leqslant ||f||_p + ||g||_p.$$

Q.E.D

6.3 Lecture 27 (More Properties on L^p spaces)

Now, we need to establish that the space is complete. We break this up into cases

Case 1: Let $p = \infty$. Let $Z_{n,m} = \{x : |f_n(x) - f_m(x)| > ||f_n - f_m||_\infty\}$. By definition, $|Z_{n,m}| = 0$. Notice that if

$$Z_{n,m} = \bigcup_{n,m} Z_{n,m},$$

then $|Z| = 0$ as well. If $x \notin Z$, then $|f_n(x) - f_m(x)| \leqslant ||f_n - f_m||_\infty$. Assume that, $\forall \epsilon > 0$, $\exists N(\epsilon)$ such that $\forall n, m \geqslant N(\epsilon)$, $||f_n - f_m||_\infty < \epsilon$. That is, $\{f_n\}$ is Cauchy in L^∞. Thus, for $x \notin Z$, we get $\{f_n(x)\}$ is also Cauchy. Since this is a sequence of real or complex numbers, the completeness of these spaces tells us that this converges. Let $f(x) := \lim_{n\to\infty} f_n(x)$. THen we need to show that $||f_n - f||_\infty \to 0$. Notice that for $x \notin Z$, $f_n(x) - f(x)| = \lim_{n\to\infty} |f_n(x) - f_m(x)| \leqslant \liminf_{n\to\infty} ||f_n - f_m||_\infty$. We certainely have that

$$\limsup_{n\to\infty} ||f_n - f||_\infty \leqslant \limsup_{n\to\infty} \liminf_{n\to\infty} ||f_n - f_m||_\infty = 0.$$

We are almost done. We need to establish that $f \in L^\infty$. We see this by noting $||f||_\infty < ||f - f_n||_\infty + ||f_n||_\infty < \infty$.

Case 2: Let $1 \leqslant p < \infty$. We follow the same general strategy. First, assume $\{f_n\}$ is Cauchy in L^p. Note by Chebychev (**Theorem 4.4**), we get that

$$|\{|f_n - f_m| > \epsilon\}| \leqslant \frac{1}{\epsilon^p}||f_n - f_m||_p^p.$$

For n, m sufficiently large, we have that $||f_n - f_m||_p^p < \epsilon^p$. So we have that $\{f_n\}$ is Cauchy in measure, which implies that $\{f_n\}$ is convergent in measure. Since it is convergent in measure, we have that there is a subsequence $\{f_{n_k}\}$ which converges to a function almost everywhere, denote this by f. The argument in Case 1 gives us that

$$||f_n - f||_p \leqslant \left(\int \lim_{k\to\infty} |f_n - f_{n_k}|^p\right)^{1/p}.$$

Fatou's Lemma (**Theorem 4.5**) then tells us

$$\left(\int \lim_{k\to\infty} |f_n - f_{n_k}|^p\right)^{1/p} \leqslant \liminf_{k\to\infty} ||f_n - f_{n_k}||_p \leqslant \epsilon$$

if $n \geqslant N(\epsilon)$. Finally, $f \in L^p$ by the same reason as in Case 1.

Q.E.D

We now want to discuss separability.

Definition. A metric space is separable if there exists a countable dense subset.

Example 6.3. $\mathbb{R}$ is separable, since $\mathbb{Q} \subseteq \mathbb{R}$ is countable and $\mathbb{Q}$ is dense.

We will see that L^p is almost always separable!

Theorem 6.10. If $1 \leqslant p < \infty$, then $L^p(E)$ is separable.

Remark 36. Why don't we have $p = \infty$ is separable? Because it's not true!

Explicitly, examine $L^\infty([0,1])$, and consider the family of function $\mathfrak{F} = \{\chi_{[0,t)}\}$. Examine

$$||\chi_{[0,t)} - \chi_{[0,s)}||_\infty = 1, \quad t \neq s.$$

Now, suppose there were a countably dense subset. Take a ball of radius 1/3 around the points in this subset. Since this is a dense set, the union of these balls would be the whole space $[0,1]$. However, each of these balls can only contain 1 function from the family, due to the fact that the L^∞ norm is 1. Therefore, we have a contradiction – the countable set must be uncountable.

Let's now prove the theorem.

Proof. Let $\mathcal{S}$ be the collection of all functions of the form

$$\sum_{i=1}^{m} q_i \chi_{Q_i},$$

where $q_i \in \mathbb{Q}$, $m < \infty$, and the Q_i are **dyadic cubes**. That is,

$$Q_k = 2^{-k}\vec{z} + [0, 2^{-k}]^n$$

for some $\vec{z} \in \mathbb{Z}$. We now have countably many functions, since our cubes are countable and the coefficients are countable. Note as well that $f \in \mathcal{S}$ implies that $f \in L^p(\mathbb{R}^n)$.

Let $\bar{\mathcal{S}}$ be the L^p closure of $\mathcal{S}$. We want to show that $\bar{\mathcal{S}} = L^p(\mathbb{R}^n)$. We break this up into steps.

Step 1. We want to show that $\chi_G \in \bar{\mathcal{S}}$ for G open with $|G| < \infty$. In such a case, a variation of **Remark 2** tells us that G is covered by dyadic cubes; that is, $G = \bigcup_i Q_i$. Then we have that

$$\chi_G = \sum_i \chi_{Q_i} \text{ a.e.}$$

Furthermore, this tells us

$$|G| = \sum_i |Q_i|.$$

Now, we look at

$$\left\| \chi_G - \sum_{i=1}^{N} \chi_{Q_i} \right\|_p = \left\| \sum_{i=N+1}^{\infty} \chi_{Q_i} \right\|_p .$$

Using the definition of the L^p norm, we get that this is

$$\left(\int \sum_{i=N+1}^{\infty} \chi_{Q_i} \right)^{1/p} .$$

Tonelli (**Theorem 4.17**) gives us that we can rearrange the sum and integral to get

$$\left(\sum_{i=N+1}^{\infty} \int \chi_{Q_i} \right)^{1/p} = \left(\sum_{i=N+1}^{\infty} |Q_i| \right) \to 0$$

as we let $N \to \infty$.

Step 2. We now want to show that $\chi_E \in \bar{\mathcal{S}}$ for E measurable, $|E| < \infty$. However, this is clear by the fact that we can estimate measurable sets with open sets.

Step 3. We now want simple functions $f \in \bar{\mathcal{S}}$ for f with finite support. This also follows by a similar argument to Step 1, though.

Step 4. We have that non-negative functions $f \in L^p$ are in $\bar{\mathcal{S}}$.

Step 5. Finally, we get that all functions $f \in L^p$ are in $\bar{\mathcal{S}}$.

Q.E.D

6.4 Lecture 28 (Missed due to OSU)

I will not include proofs here, as they will just be ripped from the book. I am mostly trying to guess what results I missed.

Theorem 6.11. If $0 < p < 1$, $L^p(E)$ is a complete, separable metric space with distance defined by

$$d(f, g) = ||f - g||_{p,E}^p.$$

Theorem 6.12. If $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$, then

$$\lim_{|h| \to 0} ||f(x + h) - f(x)||_p = 0.$$

Remark 37. This implies that continuity is preserved under L^p norms. We also remark that this theorem is true for $0 < p < 1$, however, it breaks for $p = \infty$.

Theorem 6.13. Let $1 \leqslant p \leqslant \infty$, $f \in L^p(\mathbb{R}^n)$ and $g \in L^1(\mathbb{R}^n)$. Then $f * g \in L^p(\mathbb{R}^n)$ and

$$||f * g||_p \leqslant ||f||_p ||g||_1.$$

Theorem 6.14. (Young's Convolution Theorem) Let p and q satisfy $1 \leqslant p, q \leqslant \infty$ and $1/p + 1/q \geqslant 1$, and let r be defined by $1/r = 1/p + 1/q - 1$. If $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$, then $f * g \in L^r(\mathbb{R}^n)$, and

$$||f * g||_r \leqslant ||f||_p ||g||_q.$$

6.5 Lecture 29 (Convolutions, Approx to the Identity)

We now set up some notation. If $\alpha = (\alpha_1, \ldots, \alpha_n)$, then

$$D^\alpha f(x) = \left(\frac{\partial^{\alpha_1}}{\partial x_1} \cdots \frac{\partial^{\alpha_n}}{\partial x_n} \right) f.$$

That is, this is a compact way of denoting mixed partial fractions.

Definition. $|\alpha| = \sum_{i=1}^n \alpha_i$ is defined to be the **order** of the derivative D^α.

We denoted by C^m the set of functions where $D^\alpha f$ exists and is continuous for all $|\alpha| \leqslant m$. For $m = \infty$, we have that it is the set of smooth functions. We denote by C_0^m the set of functions which are in C^m and where they have compact support.

Theorem 6.15. If $f \in L^p$ for some $1 \leqslant p \leqslant \infty$, and $g \in C_0^m$, then $f * g \in C^m$ and

$$D^\alpha (f * g) = f * D^\alpha g$$

for all $|\alpha| \leqslant m$.

Remark 38. It's enough to prove $\frac{\partial}{\partial x_i}(f * g) = f * \frac{\partial}{\partial x_i} g$, since we are just iterating the derivatives by construction.

Proof. We first want to show that $f * g$ is continuous. It is, in fact, going to be uniformly continuous. We start with examining

$$|f * g(x+h) - f * g(x)| = \left| \int f(y)\left(g(x+h-y) - g(x-y)\right) dy \right|.$$

We now do a transformation. Let $u = x - y$. Then $du = -dy$, and so we have

$$\left| \int f(x-u)\left(g(u+h) - g(u)\right) du \right| \leqslant \int |f(x-u)||g(u+h) - g(u)| du$$

We now apply Hölder! We thus have

$$\int |f(x-u)||g(u+h) - g(u)| du \leqslant ||f||_p ||g(u+h) - g(u)||_q.$$

Now, using **Theorem 5.19**, we know that we can make $||g(u+h) - g(u)||$ as small as we wish with regards to h so long as $q \neq \infty$. However, the case of $q = \infty$ follows just from the fact that g is uniformly continuous.

Now, we want to check the formula. We use **Remark 38** to note we only need to check the case of a single derivative. We start with examining

$$\frac{(f * g)(x + he_i) - (f * g)(x)}{h} = \int f(y) \left(\frac{g(x + he_i y) - g(x-y)}{h} \right) dy.$$

We can use the Mean Value Theorem to get that this is equal to

$$\int f(y) \frac{\partial}{\partial x_i} g(x - y + \xi e_i) dy$$

for some $\xi \in [0, h]$. Since $(\partial/\partial x_i)g$ is continuous with compact support, it is uniformly continuous. Therefore, we have

$$\left| \frac{\partial}{\partial x_i} g(x - y + \xi e_i) - \frac{\partial}{\partial x_i} g(x-y) \right| < \epsilon$$

if $|h| < \delta$. Note that this bound is independent of x and y. Now, examine

$$\left| \frac{f * g(x + he_i) - f * g(x)}{h} - f * \frac{\partial}{\partial x_i} g(x) \right|$$

$$= \left| \int f(y) \left(\frac{\partial}{\partial x_i} g(x - y + \xi e_i) - \frac{\partial}{\partial x_i} g(x-y) \right) dy \right|$$

$$\leqslant \int |f(y)| \left| \frac{\partial}{\partial x_i} g(x - y + \xi e_i) - \frac{\partial}{\partial x_i} g(x-y) \right| dy < \epsilon \int |f(y)| dy.$$

Now, we may use Hölder to get

$$\epsilon \int |f(y)|dy \leqslant ||f||_p \epsilon.$$

Since $f \in L^p$, we have that this goes to zero as we let $\epsilon \to 0$, as desired.

Q.E.D

Remark 39. (a) IF $g \in C_0^\infty$, $f \in L^p$, then we have that $f * g \in C^\infty$.

(b) If f does **not** have compact support, then neither does $f * g$. But, if f and g both have compact support, then $f * g$ also has compact support.

For fixed $K \in L^1$, we can define a transformation $T : L^1 \to L^1$, where $f \mapsto f * K$. We use K because K is called the **kernel** of this transformation.

Definition. For K fixed, $\epsilon > 0$, define

$$K_\epsilon(x) = \frac{1}{\epsilon^n} K\left(\frac{x}{\epsilon}\right).$$

This is called the **rescaled Kernel**.

Remark 40. (a) This is so that the integral is unchanged;

$$\int_{\mathbb{R}^n} K_\epsilon = \int_{\mathbb{R}^n} K.$$

(b) Notice that

$$\lim_{\epsilon \to 0} \int_{\{x \,:\, |x| > \delta\}} K_\epsilon(x) dx = 0.$$

That is, all the interesting information is near 0. We see that this holds true by examining the following chain of info:

$$\int_{\{x \,:\, |x| > \delta\}} |K_\epsilon(x)| dx = \frac{1}{\epsilon^n} \int_{\{x \,:\, |x| > \delta\}} \left| K\left(\frac{x}{\epsilon}\right) \right| dx.$$

We now do a transformation to get

$$\int_{\{y \,:\, |y| > \delta/\epsilon\}} |K(y)| dy.$$

We can now rewrite this as

$$\int_{\mathbb{R}^n} |K(y)| \chi_{(\delta/\epsilon, \infty)}(y) dy.$$

We want to now bring in the limit. We can use the dominated convergence theorem to do so; notice that this is dominated above by $|K(n)|$, which is integrable by assumption, and notice as well that as we let $\epsilon \to 0$, we have that the integral will be zero. Therefore, we can bring the limit in to see that this will be 0.

Why are these called approximations to the identity? Assume $\int K = 1$, so that $\int K_\epsilon = 1$. Denote by $f_\epsilon = (f * K_\epsilon)(x)$, which in other words is

$$f_\epsilon = \int f(x-y)K_\epsilon(y)dy \approx f(x).$$

We would like to study the conditions where this is actually true. That is, where the limit is actually going to be $f(x)$.

6.6 Lecture 30 (Approximations of the Identity Cont.)

Remark 41. Recall from the homework that if $K \in C_0^\infty$, then $f_\epsilon(x) \to f(x)$ a.e.

Theorem 6.16. If $\int K = 1$ and $f \in L^p$, then $f_\epsilon \to f$ in L^p. That is, $||f_\epsilon - f||_p \to 0$.

Proof. We first show pointwise convergence, and then we use this pointwise result to derive L^p convergence.

First, let's examine $|f_\epsilon(x) - f(x)|$. We may write this as

$$|f_\epsilon(x) - f(x)| = \left| \int f(x-y)K_\epsilon(y)dy - f(x) \right|.$$

Now, we use **Remark 40 (a)** to note that we can rewrite this as

$$\left| \int f(x-y)K_\epsilon(y)dy - f(x) \int K_\epsilon(y)dy \right|.$$

We can use linearity of the integrals toi simplify this to

$$\left| \int K_\epsilon(y) \left(f(x-y) - f(x) \right) dy \right| \leqslant \int |K_\epsilon(y)| \left| f(x-y) - f(x) \right| dy.$$

Now, we notice that this is *very* close to Hölder. We just need to modify the K_ϵ slightly. Let p, q be such that $1/p + 1/q = 1$. Then we write

$$K_\epsilon(y) = \left(K_\epsilon(y)\right)^{1/p} \left(K_epsilon(y)\right)^{1/q}.$$

Substituting this into the above, we get

$$\int |K_\epsilon(y)| \left| f(x-y) - f(x) \right| dy = \int |K_\epsilon(y)|^{1/p} \left| f(x-y) - f(x) \right| |K_\epsilon(y)|^{1/q} dy.$$

We use Hölder to bound this above by

$$\left(\int |f(x-y) - f(x)|^p \, |K_\epsilon(y)| \, dy \right)^{1/p} \left(\int |K_\epsilon(y)| \, dy \right)^{1/q}.$$

Now, notice that

$$\left(\int |K_\epsilon(y)|\, dy\right)^{1/q} = ||K||_1^{1/q} = ||K||_1^{1/q},$$

by **Remark 40 (b)**. So, putting this all together, we have an upper bound of

$$|f_\epsilon(x) - f(x)| \leqslant \left(\int |f(x-y) - f(x)|^p\, |K_\epsilon(y)|\, dy\right)^{1/p} ||K||_1^{1/q}.$$

Going back to what we want to show, examine $\int |f_\epsilon - f|^p = ||f_\epsilon - f||_p^p$. We want to show that this goes to 0 as ϵ goes to 0. We can use the upper bound we just derived to bound this by

$$\int |f_\epsilon(x) - f(x)|^p dx \leqslant \int\int (|f(x-y) - f(x)|^p |K_\epsilon(y)|)\, ||K||_1^{p/q} dy dx.$$

Pulling constants out and using Tonelli, we may rewrite the upper bound as

$$||K||^{p/q} \int \left(\int |f(x-y) - f(x)|^p dx\right) |K_\epsilon(y)| dy.$$

Let $g_y(x) = f(x-y) - f(x)$. Then we have that this upper bound is

$$||K||^{p/q} \int ||g_y||_p^p |K_\epsilon(y)| dy.$$

We now use **Theorem 5.19** to note that, as $|y| \to 0$, we have that $||g_y||_p^p \to 0$. Thus, for fixed ϵ', we have that there is a δ so that for $|y| < \delta$ we have $||g_y||_p^p < \epsilon'$. In the case where y is not small, i.e. $|y| \geqslant \delta$, we have that $||g_y||_p^p \leqslant 2^p ||f||_p^p$, via a homework problem. Hence, we get an upper bound of

$$||K||^{p/q} \epsilon' \left(\int_{|y|<\delta} |K_\epsilon(y)| dy\right) + (2||f||_p)^p \left(\int_{|y|\geqslant\delta} |K_\epsilon(y)| dy\right).$$

From **Remark 40 (b)**, we know that the blue (or right) portion goes to 0 as we take ϵ to 0. From **Remark 40 (a)**, we note that the red (or left) portion can be written as $||K||_1 \epsilon'$. Hence, we have

$$\limsup_{\epsilon \to 0} \int |f_\epsilon - f|^p \leqslant \epsilon' ||K||_1^{p/q+1}.$$

Since this applies for all $\epsilon' > 0$, we get that it must be 0. Hence, we have convergence in L^p norm. **Q.E.D**

Remark 42. The trick of showing pointwise convergence and then L^p convergence is a useful trick that we will do often.

Corollary 6.16.1. C_0^∞ is dense in $L^p(\mathbb{R})$ for $1 \leqslant p < \infty$.

Proof. We break this up into some cases.

Case 1. Suppose $f \in L^p$ and has compact support. Let $\int K = 1$ and $K \in C_0^\infty$. Then if we take $f_\epsilon = f * K_\epsilon$, we have a smooth function with compact support. The smoothness comes from **Theorem 5.22**, and the compact support comes **Remark 39 (a)**. Notice that this approximates f in L^p by **Theorem 5.23**, and we're done.

Case 2. Now suppose $f \in L^p$ and that it does **not** have compact support. Let $g_R = f\chi_{|x| \leqslant R}$. We claim that $g_R \to f$ in L^p. Examine

$$||g_R - f||_p = \left(\int |g_R - f|^p\right)^{1/p} = \int |f|\chi_{|x|>R}.$$

We may use the Dominated Convergence Theorem **Theorem 4.6** here. Notice that $R \to \infty$ on the inside will give us 0, and to use dominated convergence theorem we dominate this function by just $|f|$, which we know is integrable by assumption. Hence, we have that

$$\lim_{R \to 0} ||g_R - f||_p = 0.$$

By the previous case, we can approximate g_R by smooth functions with compact support, and so we have

$$||f - g * K_\epsilon||_p \leqslant ||f - g_R||_p + ||g_R - g_R * K_\epsilon|_p.$$

The things on the right go to 0 as we take $R \to \infty$ and $\epsilon \to 0$.

Q.E.D

Remark 43. How do we know C_0^∞ is nonempty? That is, how do we know there is a function which has compact support and is smooth? We will build one later on.

Definition. We say that **f(x) = O(g(x))** as $x \to x_0$ if $|f(x)/g(x)| \leqslant C$ for x near x_0. This is what is called **big O notation**.

Definition. We say that **f(x) = o(g(x))** as $x \to x_0$ if $|f(x)/g(x)| \to 0$ as $x \to x_0$. This is what is is called **little o notation**.

Using these, we can write some interesting theorems on pointwise convergence of f_ϵ which we will prove next lecture.

Theorem 6.17. If $f \in L^\infty$, then $f_\epsilon(x) \to f(x)$ at every point of continuity of f.

Theorem 6.18. If $f \in L^1$, $K \in L^1 \cap L^\infty$, and $K(x) = o(|x|^{-n})$ as $|x| \to \infty$, then $f_\epsilon(x) \to f(x)$ at every point of continuity.

Theorem 6.19. If $f \in L^1$, $K \in L^1 \cap L^\infty$, and $K(x) = O(|x|^{-n-\lambda})$ for some $\lambda > 0$, then $f_\epsilon(x) \to f(x)$ at every point in the Lebesgue set of f.

Remark 44. Note that this is for n equal to the dimension of the space L is over.

6.7 Lecture 31

We now prove the theorems. We start with **Theorem 5.24**.

Proof. We do what should now be a routine trick. Notice that we can write

$$f(x) = \int f(x)K_\epsilon(y)dy,$$

$$f_\epsilon(x) = \int f(x-y)K_\epsilon(y)dy.$$

Then we have

$$|f_\epsilon(x) - f(x)| = \left|\int K_\epsilon(y)\left(f(x-y) - f(x)\right)dy\right|$$

$$\leqslant \int |K_\epsilon(y)||f(x-y) - f(x)|dy.$$

Let $\epsilon' > 0$ be fixed. Then we pick a δ such that $|f(x-y) - f(x)| < \epsilon'$ for all $|y| < \delta$. We may do this since x is point of continuity for f. We can then break up the integral into cases; where $|y| < \delta$ and where $|y| \geqslant \delta$. That is, we have

$$\int |K_\epsilon(y)||f(x-y)-f(x)|dy = \int_{|y|<\delta} |K_\epsilon(y)||f(x-y)-f(x)|dy + \int_{|y|\geqslant\delta} |K_\epsilon(y)||f(x-y)-f(x)|dy$$

$$\leqslant \epsilon' \int_{|y|<\delta} |K_\epsilon(y)|dy + \int_{|y|\geqslant\delta} |K_\epsilon(y)||f(x-y) - f(x)|dy.$$

Now, we can utilize the proof of **Theorem 5.23** to note that, as we let $\epsilon \to 0$, we have

$$\int_{|y|\geqslant\delta} |K_\epsilon(y)||f(x-y) - f(x)|dy \to 0.$$

Hence, letting $\epsilon \to 0$, our upper bound is then

$$\epsilon' \int_{|y|<\delta} |K_\epsilon(y)|dy.$$

Now, we can extend the domain to make this bigger. That is,

$$\epsilon' \int_{|y|<\delta} |K_\epsilon(y)|dy \leqslant \epsilon' \int |K_\epsilon(y)|dy = \epsilon'||K_\epsilon||_1 = \epsilon'||K||_1.$$

So we simply have an upper bound of $\epsilon'||K||_1$ for all $\epsilon' > 0$. Taking $\epsilon' \to 0$ gives us the desired result. **Q.E.D**

We now prove **Theorem 5.25**.

Proof. We start the same way. That is, write

$$f(x) = \int f(x) K_\epsilon(y) dy,$$

$$f_\epsilon(x) = \int f(x-y) K_\epsilon(y) dy.$$

Then again we have

$$|f_\epsilon(x) - f(x)| \leqslant \int |K_\epsilon(y)||f(x-y) - f(x)| dy.$$

Now, we break it up into two integrals again; for any $\delta > 0$, we get

$$\int |K_\epsilon(y)||f(x-y)-f(x)|dy = \int_{|y|<\delta} |K_\epsilon(y)||f(x-y)-f(x)|dy + \int_{|y|\geqslant\delta} |K_\epsilon(y)||f(x-y)-f(x)|dy.$$

Use the triangle inequality to write $|f(x-y) - f(x)| \leqslant |f(x-y)| + |f(x)|$. Then we have

$$\int_{|y|<\delta} |K_\epsilon(y)||f(x-y)-f(x)|dy + \int_{|y|\geqslant\delta} |K_\epsilon(y)||f(x-y)|dy + \int_{|y|\geqslant\delta} |K_\epsilon(y)||f(x)|dy.$$

Now, fixing $\epsilon' > 0$, choose δ so that we have $|f(x-y) - f(x)| < \epsilon'$ for $|y| < \delta$. Let $\epsilon_0(\delta) = \epsilon_0$ be such that

$$\int_{|y|\geqslant\delta} K_\epsilon(y) dy < \epsilon'$$

for all $\epsilon < \epsilon_0$. Let $\epsilon_1(\delta) = \epsilon_1$ be such that

$$\sup_{|y|\geqslant\delta} |K_\epsilon(y)| < \epsilon'$$

for all $\epsilon < \epsilon_1$.

One should ask why we are able to do such a thing. We see this by noting that we can use the definition to rewrite the above as

$$\sup_{|y|\geqslant\delta} |K_\epsilon(y)| = \sup_{|y|\geqslant\delta} \frac{1}{\epsilon^n} K\left(\frac{y}{\epsilon}\right).$$

Now, since we are looking at $|y| \geqslant \delta$, we may notice that we have $|y|/\delta \geqslant 1$. In particular, we have $(|y|/\delta)^n \geqslant 1$ for all $n \geqslant 0$, and so we can use this to bound the above. That is,

$$\sup_{|y|\geqslant\delta} \frac{1}{\epsilon^n} K\left(\frac{y}{\epsilon}\right) \leqslant \frac{1}{\delta^n} \sup_{|y|\geqslant\delta} \left(\frac{|y|}{\epsilon}\right)^n K\left(\frac{y}{\epsilon}\right).$$

Now we can do a substitution on this. Let $x = y/\epsilon$. Then this can be rewritten as

$$\frac{1}{\delta^n} \sup_{|x|\geqslant\delta/\epsilon} |x|^n K(x).$$

By assumption, we have $K(x) = o(|x|^{-n})$ as $x \to \infty$. Recall that this means that

$$|K(x)||x|^n \to 0, \quad x \to \infty.$$

However, this is exactly what we have under the supremum! Therefore, taking $\epsilon \to 0$, we have that this pushes this to 0, since $x \to \infty$.

Putting all of this together and using some of what we had in the last proof, we have an upper bound of

$$|f_\epsilon(x) - f(x)| \leqslant \epsilon' ||K||_1 + \epsilon' |f(x)| + ||f||_1 \epsilon'$$

for all $\epsilon < \min\{\epsilon_0, \epsilon_1\}$. By assumption, $f \in L^1$ so that $||f||_1 < \infty$. Likewise, x is a point of continuity, and so $|f(x)| < \infty$. Since this applies for all $\epsilon' > 0$, we can take this to 0 to get the desired result. **Q.E.D**

We prove the final theorem. We will diverge from the lecture notes, as there seems to be a (very bad) typo. We will need to discuss **Riemann-Stieltjes integrals** first.

Definition. Let f and ϕ be two functions that are defined and finite on a finite interval $[a, b]$. If Γ is a partition of $[a, b]$, we arbitrarily select intermediate points $\{\xi_i\}_{i=1}^m$ satisfying $x_{i-1} \leqslant \xi_i \leqslant x_i$, and we write

$$R_\Gamma = \sum_{i=1}^m f(\xi_i)[\phi(x_i) - \phi(x_{i-1})].$$

R_Γ is called a **Riemann-Stieltjes sum** for Γ,

Definition. If $I = \lim_{|\Gamma| \to 0} R_\Gamma$ exists and is finite, then I is called the **Riemann-Stieltjes integral of f with respect to ϕ on $[a, b]$**, and is denoted by

$$I = \int_a^b f(x) d\phi(x) = \int_a^b f d\phi.$$

Remark 45. We note some nice features of this integral before moving on.

1. If $\phi(x) = x$, we just get the normal Riemann integral.
2. If f is continuous on $[a, b]$ and ϕ is continuously differentiable on $[a, b]$, then
$$\int_a^b f d\phi = \int_a^b f \phi' dx.$$
3. If ϕ is some sort of step function, it converts this into a discrete sum.
4. These integrals, for the most part, behave exactly like Riemann integrals. They are linear w.r.t to the functions. That is,
$$\int (f_1 + f_2) d\phi = \int f_1 d\phi + \int f_2 d\phi,$$
$$\int f d(\phi_1 + \phi_2) = \int_a^b f d\phi_1 + \int_a^b f d\phi_2.$$

5. If f is continuous on $[a,b]$, and ϕ is of bounded variation on $[a,b]$, then $\int_a^b f d\phi$ exists and we have

$$\left| \int_a^b f d\phi \right| \leqslant \left(\sup_{[a,b]} |f| \right) V[\phi; a, b].$$

6. We also have the mean value theorem; that is, if f continuous on $[a,b]$, ϕ bounded and increasing on $[a,b]$, then there exists a $\xi \in [a,b]$ s.t.

$$\int_a^b f d\phi = f(\xi)[\phi(b) - \phi(a)].$$

We omit the proofs for these.

We will also need to have a lemma on Riemann-Stieltjes functions.

Lemma 6.3. If f integrable over a spherical shell $a \leqslant |x| \leqslant b$ and $\phi(\rho)$ is continuous for $a \leqslant \rho \leqslant b$, $0 \leqslant a < b < \infty$. Let $F(\rho) = \int_{|a| \leqslant |x| \leqslant \rho} f(x)dx$ for $a \leqslant \rho \leqslant b$. Then

$$\int_{a \leqslant |x| \leqslant b} f(x)\phi(|x|)dx = \int_a^b \phi(\rho) dF(\rho).$$

Proof. Let $f = f^+ - f^-$. These are two bounded increasing functions, and so F is of bounded variation on $[a,b]$. Furthermore, this tells us that $\int_a^b \phi dF$ is well defined. We may assume that $f \geqslant 0$ without loss of generality. Let $I = \int_{a \leqslant |x| \leqslant b} f(x)\phi(|x|)dx$, and let $\{a = \rho_0 < \rho_1 < \cdots < \rho_k = b\}$ be a partition of $[a,b]$. Then we have

$$I = \sum_{i=1}^k \int_{\rho_{i-1} \leqslant |x| \leqslant \rho_i} f(x)\phi(|x|)dx.$$

Since $f \geqslant 0$, we get

$$\sum_{i=1}^k m_i \int_{\rho_{i-1} \leqslant |x| \leqslant \rho_i} f(x)dx \leqslant I \leqslant \sum_{i=1}^k M_i \int_{\rho_{i-1} \leqslant |x| \leqslant \rho_i} f(x)dx,$$

where M_i is the max of ϕ on $[a,b]$ and m_i is the min of ϕ on $[a,b]$. Use the fundamental theorem of calculus to rewrite this as

$$\sum_{i=1}^k m_i [F(\rho_i) - F(\rho_{i-1})] \leqslant I \leqslant \sum_{i=1}^k M_i [F(\rho_i) - F(\rho_{i-1})].$$

We squeeze this together and use **Theorem 2.24 from the book** to deduce the desired equality. **Q.E.D**

Proof. Let x_0 be a point of the Lebesgue set of f, so that

$$\rho^{-n} \int_{|x|<\rho} |f(x_0 + x) - f(x_0)| dx \to 0$$

as $\rho \to 0$. By considering the function $\widetilde{f}(x) = f(x_0 + x)$, we may assume that $x_0 = 0$. Since the hypothesis on K implies that $K(x) = o(|x|^{-n})$, the conclusion follows from the prior theorem if f is continuous at 0. Hence, subtracting from f a continuous function with compact support which equal $f(0)$ at 0, we may suppose that $f(0) = 0$.

We can now use the fact that $|K(x)|$ is bounded and that $K(x) = O(|x|^{-n-\lambda})$ to get a single estimate

$$|K(x)| \leqslant \frac{M_1}{(1 + |x|)^{n+\lambda}}.$$

Hence,

$$K_\epsilon(x) \leqslant M_1 \frac{\epsilon^\lambda}{(\epsilon + |x|)^{n+\lambda}}.$$

Therefore,

$$f_\epsilon(0)| \leqslant M_1 \int_{\mathbb{R}^n} |f(x)| \frac{\epsilon^\lambda}{(\epsilon + |x|)^{n+\lambda}} dx.$$

Now, let $F(\rho) = \int_{|x|\leqslant\rho} |f(x)| dx$. The hypotheses that $x_0 = 0$ is a Lebesgue point of f and that $f(0) = 0$ imply that given $\zeta > 0$, there is a $\delta > 0$ such that $F(\rho) < \zeta\rho^n$ if $p \leqslant \delta$. Write

$$\int_{\mathbb{R}^n} |f(x)| \frac{\epsilon^\lambda}{(\epsilon + |x|)^{n+\lambda}} dx = \int_{|x|\leqslant\delta} + \int_{|x|>\delta} = A + B.$$

Taking

$$\phi(\rho) = \frac{\epsilon^\lambda}{(\epsilon + \rho)^{n+\lambda}}$$

and $[a, b] = [0, \delta]$ in **Lemma 5.8**, we have

$$A = \int_0^\delta \frac{\epsilon^\lambda}{(\epsilon + \rho)^{n+\lambda}} dF(\rho)$$

Integrate this by parts and use $F(0) = 0$ to get

$$A = \phi(\delta)F(\delta) + (n + \lambda) \int_0^\delta F(\rho) \frac{\phi(\rho)}{(\epsilon + \rho)} d\rho.$$

The term on the left goes to 0 as $\epsilon \to 0$, and the right term, after using a transformation $\rho = \epsilon t$, has an upper bound of

$$(n + \lambda)\zeta \int_0^\delta \rho^n \frac{\phi(\rho)}{(\epsilon + \rho)} d\rho = (n + \lambda)\zeta \int_0^{\delta/\epsilon} \frac{t^n}{(1 + t)^{n+\lambda+1}} dt.$$

Hence, $\limsup_{\epsilon\to 0} A \leqslant c\zeta$ for some constant c. To estimate B, note that if $|x| > \delta$, then $\epsilon + |x| > \delta$, so that

$$B \leqslant \frac{\epsilon^\lambda}{\delta^{n+\lambda}} \int_{|x|>\delta} |f(x)|dx \leqslant \frac{\epsilon^\lambda}{\delta^{n+\lambda}} ||f||_1.$$

This goes to 0 as $\epsilon \to 0$. Taking $\zeta \to 0$ gives us our desired result. **Q.E.D**

6.8 Lecture 32 (Abstract Measure Spaces)

We briefly discuss some examples of convolution kernels before moving on.

Example 6.4. (i) The **Poison Kernel**,

$$P(x) = K(x) = \frac{1}{\pi(1+x^2)}, x \in \mathbb{R}.$$

(ii) The **Gauss-Weierstrauss** Kernel,

$$K(x) = \frac{1}{\sqrt{\pi}} e^{-x^2}, x \in \mathbb{R}.$$

Remark 46. We can generalize this to higher dimensions as well.

(iii) The **Fejer Kernel**,

$$K(x) = \frac{1}{\pi}\left(\frac{\sin(x)}{x}\right)^2, x \in \mathbb{R}.$$

There are applications of using these to sort of normalize or control continuous functions, but I will skip over these for conciseness.

Chapter 7

Abstract Measure Spaces

Definition. An **(abstract) measure space** is a triple $(X, \mathscr{F}, \mu)$, where X denotes the space, $\mathscr{F}$ denotes a collection of subsets, and μ is a measure, which satisfies the following properties:

(i) $\mathscr{F}$ is a σ-algebra of subsets of X,

(ii) μ is a measure on $\mathscr{F}$:

(a) $\mu : \mathscr{F} \to [0, \infty]$ is a set function,

(b)
$$\mu\left(\bigcup_k E_k\right) = \sum_k \mu(E_k)$$
if $E_k \in \mathscr{F}$ and $E_k \cap E_j = \varnothing$ if $j \neq k$.

Example 7.1. (i) Let $X = \mathbb{R}^n$, $\mathscr{F}$ the collection of Lebesgue measurable sets (or Borel measurable sets), and suppose there is a fxed non-negative measurable set function f. Define $\mu(A) = \int_A f(x)dx$. Then we have that μ satisfies the properties from the definition.

(ii) Let $X = \mathbb{Z}$ (or some countable set), $\mathscr{F} = \mathcal{P}(X)$ (the powerset of X), and define $\mu(A) = |A|$ (that is, the cardinality of A). This is called the **countable measure**.

(iii) Let $X = \mathbb{Z}$, $\mathscr{F} = \mathcal{P}(X)$, and fix a non-negative sequence $\{a_k\}_{k\in\mathbb{Z}}$. Set $\mu(\{k\}) = a_k$. Then $\mu(A) = \sum_{k\in A} a_k$. This gives us a **probability measure** if $\sum a_k = 1$.

We now note some properties.

Lemma 7.1. If $(X, \mathscr{F}, \mu)$ is a measure space, we have the following:

(i) $\mu(\varnothing) = 0$ as long as $\mu(A) < \infty$ for some A.

(ii) If $A \subseteq B$, then $\mu(A) \leqslant \mu(B)$.

(iii) For any collection $E_k \in \mathscr{F}$, we have

$$\mu\left(\bigcup_k E_k\right) \leqslant \sum_k \mu(E_k).$$

(iv) If $E_k \nearrow E$, then $\mu(E_K) \to \mu(E)$. Likewise, if $E_k \searrow E$, and $\mu(E_{k_0}) < \infty$ for some k_0, then $\mu(E_k) \to \mu(E)$.

Proof. (i) If $\mu(A) < \infty$, then we have that $\mu(A \cup \varnothing) = \mu(A) + \mu(\varnothing) = \mu(A)$. Subtracting $\mu(A)$ from both sides gives $\mu(\varnothing) = \varnothing$.

(ii) Write B as $B = (B - A) \cup A$. Then these are disjoint, and so we have $\mu(B) = \mu(B - A) + \mu(A)$. Therefore, we have $\mu(A) \leqslant \mu(B)$.

(iii) Write $F_1 = E_1$, $F_k = E_k - \left(\bigcup_{i=1}^{k-1} E_i\right)$. Then $\bigcup_k F_k = \bigcup_k E_k$, and furthermore $F_k \cap F_j = \varnothing$ if $k \neq j$. So we can write this as

$$\mu\left(\bigcup_k F_k\right) = \sum_k \mu(F_k).$$

Now, $\mu(F_k) \leqslant \mu(E_k)$ for all k, and so we have

$$\mu\left(\bigcup_k E_k\right) = \left(\bigcup_k F_k\right) \leqslant \sum_k \mu(E_k).$$

(iv) This is analogous to the proof of **Proposition 2.1 (ii)**.

Q.E.D

We can use this to define abstract measurable functions.

Definition. Let $(X, \mathscr{F})$ and $(Y, \mathscr{G})$ be two spaces and σ-algebras. Suppose we have a function $f : X \to Y$. Then we say that f is a **measurable function** if $f^{-1}(B) \in \mathscr{F}$ for all $B \in \mathscr{G}$.

So, in this language, what are the measurable functions that we have been using? They are functions $f : (\mathbb{R}^n, \mathscr{M}) \to (\mathbb{R}, \mathscr{B})$, where $\mathscr{M}$ is the set of all Lebesgue measurable sets and $\mathscr{B}$ is the set of all Borel measurable sets. We've been only using open sets, not Lebesgue measurable sets!

7.1 Lecture 33

We go over some properties of $\mathcal{F}$-measurable sets.

Lemma 7.2. (i) If f and g are $\mathcal{F}$-measurable, then so are $f + g$ and fg. If $c \in \mathbb{R}$, then cf is $\mathcal{F}$-measurable. If ϕ is continuous, then $\phi \circ f$ is $\mathcal{F}$-measurable.

(ii) If $\{f_n\}_n$ is a sequence of $\mathcal{F}$-measurable functions, then the following are also $\mathcal{F}$-measurable:

(a) $\sup_n f_n$,

(b) $\inf_n f_n$,

(c) $\limsup_n f_n$,

(d) $\liminf_n f_n$,

(e) if the limit exists, $\lim_n f_n$.

(iii) If f is non-negative and $\mathcal{F}$-measurable, then there exists non-negative simple functions which are $\mathcal{F}$-measurable such that $f_k \nearrow f$.

The proof to these is analogous to the proofs found in **Chapter 2**. We remark here that Durett ([2]) assumes that μ is σ-finite. We define this below.

Definition. A measure μ is said to be **σ-finite** if there exists a sequence $E_n \subseteq X$ such that $E_n \nearrow X$ and $\mu(E_n) < \infty$ for all n.

We will not use this assumption, however.

Definition. We define the **(abstract) integral for a simple $\mathcal{F}$-measurable function f which is non-negative** to be

$$\int f d\mu = \sum_{i=1}^{n} a_i \mu(E_i),$$

where

$$f = \sum_{i=1}^{n} a_i \chi_{E_i},$$

and $a_i \geqslant 0, E_i \in \mathcal{F}, n < \infty$. We also define $\infty \cdot 0 = 0 \cdot \infty = 0$ in this.

We now list some properties of this new integral.

Lemma 7.3. Throughout, let $f, g \geqslant 0$ be $\mathcal{F}$-measurable simple functions. We then have the following:

(i)

$$\int (af) d\mu = a \int f d\mu,$$

(ii)

$$\int (f+g) d\mu = \int f d\mu + \int g d\mu,$$

(iii) If $E \subseteq F$, then $\int_E f d\mu \leqslant \int_F f d\mu$.

(iv) If $f \leqslant g$, then $\int f d\mu \leqslant \int g d\mu$.

Proof. (i) Examine $\int (af)d\mu$. Since f is a simple function, we have that af is defined to be

$$af = \sum_{i=1}^{n} a a_i \chi_{E_i}.$$

Hence,

$$\int (af)d\mu = \sum_{i=1}^{n} a a_i \mu(E_i) = a \sum_{i=1}^{n} a_i \mu(E_i) = a \int f d\mu.$$

(ii) This is a little more tricky. Let f be as in the definition and let g be defined by

$$g = \sum_{j=1}^{m} b_j \chi_{F_j}.$$

Then we write $f + g$ as

$$f + g = \sum_{i,j=1} (a_i + b_j)\chi_{E_i \cap F_j}.$$

Now, we can write the integral as

$$\int (f+g)d\mu = \sum_{i,j=1} (a_i + b_j)\mu(E_i \cap F_j).$$

Distribute the a_i and b_j to get this in the form

$$\int (f+g)d\mu = \sum_{i,j=1} a_i \mu(E_i \cap F_j) + \sum_{i,j=1} b_j \mu(E_i \cap F_j).$$

Now, notice that since the F_j are disjoint and their union is the whole space (and likewise with the E_i), we get that this is

$$\int (f+g)d\mu = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i \mu(E_i \cap F_j) + \sum_{j=1}^{m}\sum_{i=1}^{n} b_j \mu(E_i \cap F_j) = \sum_{i=1}^{n} a_i \mu(E_i) + \sum_{j=1}^{m} b_j \mu(F_j).$$

(iii) We first need a definition.

Definition. We define the **(abstract) integral over E of f** to be

$$\int_E f d\mu = \int f \chi_E d\mu.$$

Now, using this, we have that this result is clear. That is, we have

$$\int_E f d\mu = \int f \chi_E d\mu = \sum_{i=1}^{n} a_i \chi_E \chi_{E_i} \leqslant \sum_{i=1}^{n} a_i \chi_F \chi_{E_i} = \int_F f d\mu.$$

(iv) This is clear.

Q.E.D

We get an analogous definition to almost everywhere with regards to the measure.

Definition. We say that a property holds **μ-a.e.** if the set of points where it does not hold has μ-measure 0.

Example 7.2. We have $f \leqslant g$ μ-a.e. if $\mu(\{f > g\}) = 0$.

Using this, we can rewrite property (iv) above as follows:

Lemma 7.4. If $f \leqslant g$ μ-a.e., then $\int f d\mu \leqslant \int g d\mu$.

We now want to start to expand the integral to general non-negative functions (and then to general functions using the same procedure as prior).

Definition. We have the **(abstract) integral of a non-negative $\mathcal{F}$-measurable function f** is defined to be

$$\int f d\mu = \sup_{\substack{0 \leqslant g \leqslant f \\ g \text{ simple}}} \int g d\mu.$$

We also define the integral over a set E analogously; that is, if $E \in \mathcal{F}$, then

$$\int_E f d\mu = \int f \chi_E d\mu.$$

We now get some properties on this integral.

Lemma 7.5. (i)

$$\int (af) d\mu = a \int f d\mu.$$

(ii) If $f \leqslant g$ μ-a.e., then

$$\int f d\mu \leqslant \int g d\mu.$$

(iii) We have Chebychev's Inequality; that is,

$$\mu(\{f \geqslant a\}) \leqslant \frac{1}{a} \int f d\mu.$$

Note that we do not have additivity yet. This is because we will need the monotone convergence theorem to prove this.

Proof. (a) By definition,

$$\int (af) d\mu = \sup_{\substack{0 \leqslant g \leqslant af \\ g \text{ simple}}} \int g d\mu.$$

However, this is equivalent to

$$\sup_{\substack{0 \leqslant g \leqslant f \\ g \text{ simple}}} \int ag d\mu.$$

Now, we can use properties of simple non-negative measurable functions to rewrite this as

$$a \sup_{\substack{0 \leqslant g \leqslant f \\ g \text{ simple}}} \int g d\mu = a \int f d\mu.$$

(b) This is clear. If $f \leqslant g$, then all simple functions $h \leqslant f$ have the property that $h \leqslant g$ as well. Use this to get the inequality.

(c) Notice that we can write this as

$$\int f d\mu \geqslant \int a\chi_{\{f \geqslant a\}} d\mu,$$

and then we can note that

$$a \int \chi_{\{f \geqslant a\}} d\mu = a\mu(\{f \geqslant a\}).$$

Q.E.D

Theorem 7.1. (Fatou's Lemma) Assume $f_n \geqslant 0$ $\mathcal{F}$-measurable, then

$$\int \left(\liminf_{n\to\infty} f_n \right) d\mu \leqslant \liminf_{n\to\infty} \int f_n d\mu.$$

Proof. Let $g_n = \inf_{k \geqslant n} f_k$. Notice that these g_k are increasing, and furthermore they are increasing to $\liminf_{n\to\infty} f_n = g$. On the other hand, we have

$$\int g_n d\mu \leqslant \int f_n d\mu,$$

since g_n is an infimum over functions including f_n. So, taking the liminf of both sides, we have

$$\lim_{n\to\infty} \int g_n d\mu \leqslant \liminf_{n\to\infty} \int f_n d\mu.$$

So it suffices to show that

$$\int g d\mu \leqslant \lim_{n\to\infty} \int g_n d\mu.$$

We now use the definition; that is, for h simple,

$$\int g d\mu = \sup_{0 \leqslant h \leqslant g} \int h d\mu.$$

We need to show that if $h \leqslant g$ simple, then

$$\int h d\mu \leqslant \lim_{n\to\infty} \int g_n d\mu,$$

since if we can do this for all such h we can use the least upper bound property of the supremum. We break it up into three cases.

Case 1: Suppose $h = 0$ μ-a.e. Then we win by default.

Case 2: Now suppose

$$0 < \int h d\mu < \infty.$$

Let $E = \{x \ : \ 0 < h(x) < \infty\}$. Then we have

$$\int h d\mu = \int_E h d\mu.$$

We claim that $\mu(E) < \infty$. This is because h is simple;

$$\int h d\mu = \sum a_i \mu(E_i),$$

and so since $h < \infty$ μ-a.e. we must have that $\mu(E) < \infty$. Now, fix $\epsilon > 0$, and for all n set $E_n = \{x \in E \ : \ g_n(x) > (1-\epsilon)h(x)\}$. We have $g_n \nearrow g \geqslant h$, so eventually every point in E will be in one of these E_n. We therefore get $E_n \nearrow E$. So $(E - E_n) \searrow \varnothing$, and therefore $\mu(E - E_n) \to 0$ (since $\mu(E) < \infty$).

Now, examine

$$\int h d\mu = \int_E h d\mu = \int_{E-E_n} h d\mu + \int_{E_n} h d\mu.$$

On the left, we have that this is less than or equal to $||h||_\infty \mu(E - E_n)$, which we know goes to 0 as $n \to \infty$. On the right, we have that this is less than or equal to

$$\frac{1}{1-\epsilon} \int_{E_n} g_n d\mu \leqslant \frac{1}{1-\epsilon} \int g_n d\mu.$$

So therefore, we have that

$$\int h d\mu < \frac{1}{1-\epsilon} \int g_n d\mu$$

for n sufficiently large and for all $\epsilon > 0$. Taking the limit as ϵ goes to 0 then gives us

$$\int h d\mu \leqslant \int g_n d\mu.$$

Case 3: Now assume that $\int h d\mu = \infty$. Since h is a simple function, we must have a set $A \subseteq X$ with $\mu(A) = \infty$, $h(x) = a$ on A. Let $A_n = \{x \in A : g_n > a/2\}$. Since $g_n \to g \geqslant h = a$ on A, then $A_n \nearrow A$. Therefore, $\mu(A_n) \nearrow \mu(A) = \infty$, and so we have

$$\int g_n d\mu \geqslant \int_{A_n} g_n d\mu \geqslant \frac{a}{2}\mu(A_n) \to \infty$$

as $n \to \infty$. Hence, we have that $\int h d\mu \leqslant \lim_{n\to\infty} \int g_n d\mu$.

Q.E.D

Index

Bibliography

[1] Richard Wheeden, Antoni Zygmund. *Measure and Integral*. CRC Press, 2015.

[2] Richard Durrett. *Probability: Theory and Examples* Cambridge Series in Statistical and Probabilistic Mathematics, 2010.

[3] H.L. Royden, P.M. Fitzpatrick. *Real Analysis* Pearson Education, 2010.

[4] David McReynolds. *Math 341: Real Analysis.* *https://drive.google.com/file/d/0B26r67X0Oet3bzRWbVB3N1FFYUk/view*

[5] E Kowalski. *Measure and Integral (December 2011).* *https://people.math.ethz.ch/ kowalski/measure-integral.pdf*