M381C NOTES

ARUN DEBRAY
SEPTEMBER 14, 2015

These notes were taken in UT Austin's Math 381c class in Fall 2015, taught by Luis Caffarelli. I live-TEXed them using `vim`, and as such there may be typos; please send questions, comments, complaints, and corrections to a.debray@math.utexas.edu.

Contents

Lecture 1.

Outer Measure and the Lesbegue Measure: 8/26/15

The book for the class is *Measure and Integral*, 2^{nd} *Edition*, by Wheeden and Zygmund. This course will cover Lesbegue integration (chapters 3 to 6 of the book), function spaces (including L^p spaces; chapters 7 to 9), abstract integration and measure theory (chapters 2, 10, and 11).

In analysis, we first started with $\mathbb{R}^n$, and then started discussing functions not just as isolated entities, but as elements of a space, discussing the theory of continuous functions as a whole. For example, one might consider $C[0,1]$, the space of continuous functions on the interval $[0,1]$; then, you can talk about the distance between functions, e.g. $d(f,g) = \sup|f-g|$. For example, one can take an interval of size h around an f, so any g with $d(f,g) \le h$ is always trapped within that band. This distance is used to discuss uniform convergence: $\{f_n\}$ is said to *converge uniformly* to f if $d(f_k, f) \to 0$: that is, no matter how small you make the strip around f, for sufficiently large k, f_k is trapped in the strip.

On $\mathbb{R}^n$, the distance function is

$$d(x,y) = \left(\sum (x_i - y_i)^2\right)^{1/2},$$

but there are other distances, e.g.

$$d_1(x,y) = \sum |x_i - y_i| \qquad \text{and} \qquad d_\infty(x,y) = \sup_{1\le i\le n} |x_i - y_i|.$$

These are more or less the same, but for discussing distances between functions they do matter. For example, one distance between functions f and g is given by

$$d(f,g) = \int_0^1 |f-g|\,\mathrm{d}x.$$

But this means that there are sequences of continuous functions that converge to discontinuous ones, so $C[0,1]$ isn't the right space to study: it's not complete, just like the rational numbers. Analysis is about limits, so we really should use a complete space.

All right, great, so let's just use Riemann-integrable functions. This isn't sufficient either: let $\varepsilon > 0$ and let B_k be the function that traces out a triangle with base $\varepsilon/4^k$ wide and height 2^k. These are all nice, continuous and integrable functions, and the integral is always less than ε. Taking the limit should produce a function with integral zero, but choosing any partition doesn't work out, and so the Riemann integral isn't right either.

Let's pass from integration to defining the volume of a set. This is linked to integration, because the integral of the *indicator function* of a set S (that is, the function 1 on S and 0 outside). If S has a nice boundary, then the area of S is the Riemann integral of the indicator function. But if S doesn't have a nice boundary, the upper and lower Riemann sums disagree: for a trivial example, take $S = \mathbb{Q}$.

This is because we don't measure sets in a sharp enough way here; if you're allowed to take infinite partitions, we can cover each rational number with a cube of size $\varepsilon/2^k$, and this covers all of $\mathbb{Q}$ with total measure at most ε.

Given a theory of Lesbegue integration, you can define the measure of a set from the Lesbegue integral of its characteristic function. Going in the other direction, though, given a function f to integrate, add one more dimension and stretch a distance of 1 in that direction; then, the Lesbegue integral of f is the same as the Lesbegue measure of the new undergraph. So the integral and the measure are the same.

The Riemann integral greatly depends on the smoothness of the function; Lesbegue integration relies on a different idea. Given the graph of a function, we can draw horizontal bars across its undergraph, and calculate the area that way. Then, make finer and finer partitions, and take the limit. This is nice because it's monotone, an approximation below, and so it has nice convergence properties. But this means we need a very sturdy theory of measuring sets, since the undergraph could have a very complicated boundary. Specifically, we need a theory that allows us to do infinite processes.

The property we need is *countable additivity*: the sum of the measures of countably many disjoint sets is the measure of their union. Riemann integration has that property for a finite number of sets. An equivalent way to think of this is, given a family of monotonically increasing sets, the measure of the limit is the limit of the measures for the Lesbegue measure.

We can also think of this in terms of cubes covering a set S; given a countable number of cubes Q_j such that $\bigcup_{j=1}^{\infty} Q_j \supset S$, then we know that $\sum V(Q_j)$ is at least the measure of S, whatever that is. And we know $V(Q_j)$ is the product of its sides.

But this might have too much measure in it, so let's take the infimum over all covers.

The first observation we need is that we may refine the partition so that the cubes are nonoverlapping (i.e. their interiors are disjoint). Another observation is that closed and open covers are equivalent here; for any closed cover $\{Q_j\}$ we can choose a cover of open cubes $\{Q_j^*\}$, where $V(Q_j^*) \le V(Q_j) + \varepsilon/2^j$; thus, the sum of the measures of the open cover is at most ε plus that of the closed cubes; over all measures, the infima are the same.

Definition. This is called the *exterior measure* of S, or $|S|_e$, the infimum over all such covers by a countable nunber of nonoverlapping cubes.

One can refine a partition in a process called *dyadic refinement*: given a partition, split each cube into 2^n cubes by cutting down the middle of each side, and then throwing out all of the cubes that are disjoint with S. In fact, given a standard grid of length 1 of $\mathbb{R}^n$, repeatedly doing dyadic refinement makes for a cover of S (the *dyadic cover*) that, when S is open, has measure equal to the exterior measure of S. That is:

Lemma 1.1. *If $\{Q_j\}$ is the dyadic cover of an open set S, then $|S|_e = \sum \mathrm{Vol}(Q_j)$.*

Proof. First, notice that $Q^{(m)} = \bigcup_{i=1}^{m} Q_j$ is a compact set, so suppose we have any other cover $\{R_j\}$ of S. Then, without loss of generality, by adding $\varepsilon/2^k$, we can make it an open cover. Since $Q^{(m)}$ is compact, a finite number of the R_j cover it. Thus, the volume of this cover must be larger than that of $Q^{(m)}$, since we have a nice finite number of cubes. ⊠

Thus, the exterior measure is realized by this partition; there's no extra. It's completely tight.

To recap:

(1) We defined the exterior measure $|E|_e = \inf \sum V(Q_j)$ over all covers Q_j.
(2) We can use either open or closed cubes.
(3) For an open set S, the exterior measure of S is realized by nonoverlapping cubes.

Thus, for any set E, we have $|E|_e = \inf |U|_e$, over all open sets $U \supseteq E$. Why? For any cover by cubes of E, we can enlarge them to make it open, so the union of the cubes will be an open set, which is a superset of E. But we can choose such an open cover by cubes with measure at most ε more than $|E|_e$, for any $\varepsilon > 0$. Of course, we will want open sets to be measurable sets.

We also want this to be invariant under change of coordinates; suppose $\{Q_j\}$ is one system of coordinates and $\{Q_j^*\}$ is another system of coordinates, then the exterior measures induced by them coincide, because a cube in $\{Q_j\}$ can be approximated arbitrarily well by a countable cover in $\{Q_j^*\}$.

Some of this may feel unrigorous; careful reasoning, involving some epsilons, should fix this.

So far we've just done coverings and counting; now we come to the decision making.

Definition.

- For an open set U, we define the *Lesbegue measure* of U to be $|U|_e$, its exterior measure.
- A set E is *measurable* if for all $\varepsilon > 0$, there exists an open $U \supseteq E$ such that $|U \setminus E|_e < \varepsilon$.

The idea behind measurability is that a measurable set should look a lot like an open set. Specifically, its boundary should be well-behaved: it looks like the boundary of an open set.

Nonmeasurable sets exist, as long as you're willing to invoke the Axiom of Choice. You might be used to counterexamples like the Cantor set, but that's beautiful compared to a typical nonmeasurable set. For each $x \in [0,1)$, consider the set $x + \mathbb{Q}$, which is the equivalence class of x where $x \sim y$ if $x - y \in \mathbb{Q}$. Then, construct a set S by choosing one point from each equivalence class (which requires the Axiom of Choice). Thus, for any $q \in \mathbb{Q}$, $q + S$ is disjoint from S. In particular, $\{q + S \mid q \in \mathbb{Q}\}$ is a countable disjoint family of sets whose union is $[0,1]$. Since they're all translations, they all have the same measure, and by countable additivity, the measure of their union is 1. But there's no number a such that $\sum_1^\infty a = 1$. Thus, S is nonmeasurable.

Lecture 2.

Defining the Lesbegue Measure: 8/31/15

"And since we're analysts, or, well, at least in an analysis course..."

Recall that we started off by talking about the distance between functions, with the goal of turning the space $C[0,1]$ of continuous functions on the unit interval into a Banach space. It's already a vector space, and with the uniform distance $d(f,g) = \sup|f-g|$ induced from the uniform norm, it's in fact complete. But under the L^1 norm

$$d(f,g) = \int_0^1 |f-g|\,\mathrm{d}x,$$

this space isn't complete: limits of some Cauchy sequences of continuous functions aren't continuous. This is the first difficulty that arises from working with the space of continuous functions.

More worryingly, the space of Riemann-integrable functions isn't complete in this norm either: you can produce a Cauchy sequence of continuous functions that don't converge, e.g.

$$f_k = \sum_{i=1}^k \chi_{(q_i - \varepsilon 2^{-i}, q_i + \varepsilon 2^{-i})},$$

where q_i is an enumeration of the rationals on $[0,1]$. Any interval contains a rational, so the upper Riemann sum of the limit is always 1, but the limit of the integrals will be ε.

Since integrating indicator functions comes from defining the volumes (or measures) of their indicated sets, maybe we should step back and look at measures of sets. In the above case, the union of the intervals that defined the f_k has upper Riemann measure equal to 1 and lower measure ε; the problem is that Riemann measure is essentially finite; it can only consider finite coverings, which are too imprecise to handle the f_k. Thus, Lesbegue measure generalizes this to countable coverings, which are good enough to make this work.

Measure and integral are closely tied: measure is the integral of a characteristic set, and the integral can be calculated by approximating a function with a sum of step functions.

For Lesbegue measure, let's for now talk only about bounded sets (dealing with the unbounded case is a minor modification). Let $E \subset \mathbb{R}^n$; then, for any countable covering $\{Q_j\}$, we take $\sum \mathrm{Vol}(Q_j)$ to be an upper bound for the measure of E. Thus, the infimum of all such measures is called the *exterior measure* of E, denoted $|E|_e$.

The first issue is that not every set is measurable. So we'll just restrict to the measurable sets, but we want to define a measure on as many sets as possible: open sets, closed sets, and countable unions thereof.

We have the following easy fact.

Claim. For any measurable E and F, $|E \cup F|_e \le |E|_e + |F|_e$.

Proof. If $\{Q_j\}$ is a countable cover of E and $\{Q_j'\}$ is one for F, then their union is a cover for $E \cup F$. ⊠

Last time, we also talked about the dyadic cover: given an open set U, we build a non-overlapping covering that's inside U, by dividing it into a grid and then subdividing boxes that are partially outside of U, and so on, which is a countable process. Thus, for open sets, the exterior measure can be calculated with interior subsets, so in some sense it's also an interior measure.

What Lesbegue did was to declare open sets measurable; they're the base of the σ-algebra of measurable sets. That is, if U is open, we define the measure of U to be its exterior measure: $|U| = |U|_e$. Thus, for any E, $|E|_e = \inf_{U \supset E} |U|_e$, for U open, because any covering can be arbitrarily well approximated with an open covering.

So, to recap:

- We have defined the exterior measure $|E|_e$.
- $|E \cup F|_e \le |E|_e + |F|_e$.
- $|E|_e = \inf |U|_e$ for open $U \supset E$.

Using the last point, for each $k \in \mathbb{N}$, choose a U_k such that $|U_k|_e \le |E|_e + 1/k$. Then,

$$|E|_e = \left| \bigcap U_k \right|_e,$$

and $H = \bigcap U_k$ is a G_δ set, as it's a countable intersection of open sets. Note that H itself might not be open, but the takeaway here is that G_δ sets ought to be measurable.

The problem with Riemann measure was that the boundary of E might be too complicated for Riemann approximation to work. So we see that we should look at the boundary to determine whether a set is measurable. One could approach this by declaring G_δ sets to be measurable and then defining E as above to be measurable if $|H \setminus E|_e = 0$, but it'll be better to define things differently, and recover this as a theorem.

Definition.

- A set E is *measurable* if for all $\varepsilon > 0$, there exists an open $U \supset E$ such that $|U \setminus E|_e < \varepsilon$.
- If $|E|_e = 0$, then E is therefore measurable, and is said to have *measure zero*.

Another characterization is that E is measurable if for any $\varepsilon > 0$ there exists a measurable set H such that $|E \bigtriangleup H|_e < \varepsilon$.[1]

Armed with these definitions, we can prove useful properties of the Lesbegue measure.

Definition. A σ-algebra X is a family of sets closed under countable operations: that is, if $\{U_i\}_{i=1}^{\infty} \in X$, then $\bigcap U_i$ and $\bigcup U_i$ must be in X, as well as $\mathcal{C}(U_i)$ (that is, its complement) for each i.

Theorem 2.1. *The family of Lesbegue-measurable sets is a σ-algebra, and the Lesbegue measure is countably additive.*

We'll prove this in a few smaller steps.

Definition. The *distance* between two sets A and B is $d(A,B) = \inf_{x \in A, y \in B} d(x,y)$.

Proposition 2.2. *Suppose A and B have positive distance (that is, $d(A,B) > 0$). Then, $|A \cup B|_e = |A|_e + |B|_e$.*

Proof. The idea is to find coverings of A and B that are disjoint. Let $\delta = d(A,B)$. Then, take a grid of length $\delta/2$ to cover A and B; then, every grid box covers A xor covers B. Let U_1 be the union of those (open) boxes that cover A, and U_2 be that for B.

For any covering $\mathcal{Q}$ of $A \cup B$, we can intersect it with U_1 and U_2 to get a covering $\mathcal{Q}'$ of $A \cup B$ that is at most as large as the original one. TODO I got really lost here. ⊠

In particular, since nonintersecting compact sets have positive distance, we can apply this to compact sets.

Proposition 2.3. *Compact sets are measurable.*

[1] Here, $\bigtriangleup$ denotes symmetric difference.

Proof. Let K be compact and $U \supset K$ be open such that $|K|_e \geq |U|_e - \varepsilon$. Then, we want to show that $|U \setminus K|$ is small. Well,

$$|U \setminus K| = \sum_1^\infty \mathrm{Vol}(Q_j)$$

for a countable set of cubes Q_j (since $U \setminus K$ is open), but if we replace the infinite sum with the partial sum $\sum_1^M \mathrm{Vol}(Q_j)$ for a sufficiently large M, we have that

$$\sum_1^M \mathrm{Vol}(Q_j) = |U \setminus K| - \varepsilon.$$

But this finite union of the closed cubes Q_j is a compact set. We'll call it L_M. Therefore, $|K \cup L_M|_e = |K|_e + |L_M|_e$, since $K \cap L_M = \emptyset$, and since $K \cup L_M \subset U$, then $|K \cup L_M|_e \leq |U|_e = |K|_e + \varepsilon$, so this means that $|L_M|_e < \varepsilon$. If we let $M \to \infty$, the measure of L_M is always less than ε, i.e. the measure of $U \setminus K$, which is the union of the sets L_M as $M \to \infty$, is also less than ε, and therefore K is measurable. ⊠

In more words, the idea is that we can surround a compact set by an open set arbitrarily well, and their difference is a strip around the compact set, whose measure can be approximated by a bunch of cubes. The key idea is that, if we take only finitely many of these cubes, they're compact, so the exterior measure is additive.

Corollary 2.4.

(1) *If U is open and $K \subset U$ is compact, then $|U| = |U \setminus K| + |K|$.*
(2) *If E is any measurable set, then for any $\varepsilon > 0$ there exists a compact $K_\varepsilon \subset E$ such that $|E| \leq |K_\varepsilon| + \varepsilon$.*

Proof. The first part comes directly from the proof; for the second part, given a measurable E and $\varepsilon > 0$, there exists an open $U \supset E$ such that $|U \setminus E| < \varepsilon$. Then, since U is open, we can choose a compact $K \subset E \subset U$ such that $|U \setminus K| < \varepsilon$, which therefore also controls $|U \setminus E|$. ⊠

Some abstract approaches to measure theory start by defining exterior and interior measures, with results like that one.

Now that we have compact sets, we can start talking countable additivity.

Proof of Theorem 2.1. Corollary 2.4 means that the complement of a measurable set is measurable; it says that if a set is well-approximated by an open set, its complement is at least as well approximated by an open set.

For countable additivity, let E_j be a countable collection of disjoint measurable sets and $E = \bigcup_j E_j$. We want to show that

$$|E| = \sum_j |E_j|.$$

First, if there are finitely many E_j, we know that

$$\left|\bigcup_{j=1}^N E_j\right| \leq \sum_{j=1}^N |E_j|,$$

but we can also approximate with compact sets: let $K_j \subset E_j$ with $|K_j| \geq |E_j| - \varepsilon 2^{-j}$. Since the measure is additive on compact sets, then

$$\left|\bigcup_{j=1}^N K_j\right| = \sum_{j=1}^N |K_j| \geq \sum |E_j| - \varepsilon,$$

and we already know $\sum |K_j| \leq \sum |E_j|$ since $K_j \subset E_j$. Then, for infinitely many sets, the same argument works, but take the limit. ⊠

Now that we know that the set of Lesbegue-measurable sets is a σ-algebra, let's try to characterize it.

Definition. The *Borel sets* are the elements of the σ-algebra generated by all open sets (i.e. the smallest σ-algebra containing all open sets, or the collection of countable unions, countable intersections, and complements generated by the open sets).

However, these aren't the Lesbegue-measurable sets! That is, there exist Lesbegue-measurable sets that aren't Borel. Let's consider the *Cantor set*, which is an uncountable set of measure zero. It's closed, and therefore Borel, but will be an important step in the construction.

The Cantor set can be constructed in the following way: start with $[0,1]$. In the first step, remove the middle third; in the next step, remove the middle third of the remaining two pieces, and so on, removing the middle third of each piece at each step.

This is a countable intersection of closed sets, and thus closed, and thus measurable. But at the k^{th} step the remaining measure is $(2/3)^k$, and this goes to 0, so the Cantor set has measure zero.

One can characterize the Cantor set as the numbers

$$\sum_{k=1}^{\infty} a_k 3^{-k}, \qquad a_k = 0 \text{ or } 2.$$

But this is uncountable, because for any $x \in [0,1]$, x has a base-2 decimal expansion of 0s and 1s; replace every 1 with a 2 to get a unique point in the Cantor set. Thus, this skinny set has the same cardinality as $\mathbb{R}$!

In particular, the set of subsets of the Cantor set has cardinality $2^{\mathbb{R}}$, but there are "only" uncountably many Borel sets, so there must exist a subset of the Cantor set that isn't Borel.

Lecture 3.

Measurable Functions: 9/2/15

"It is frequently claimed that Lebesgue integration is as easy to teach as Riemann integration. This is probably true, but I have yet to be convinced that it is as easy to learn." – Thomas Körner

Recall that we've proved that the Lesbegue measure can be approximated arbitrarily well as an outer measure by open sets, and therefore (using complements) as an inner beasure using compact sets. We also proved countable additivity: if A_j is a sequence of disjoint measurable sets, then $\left|\bigcup A_j\right| = \sum |A_j|$. Using compact sets, we can show $\le$: cover A_j with a countable collection of compact sets (e.g. cubes) Q_j^k; then, $|A_j| \le \sum \operatorname{Vol}(Q_j^k)$ and

$$\left|\bigcup_{j,k} Q_j^k\right| \ge \left|\bigcup_j A_j\right|.$$

Specifically, it's possible to choose Q_j^k such that $\sum_k \operatorname{Vol}(Q_j^k) \le |A_j| + \varepsilon/2^j$; thus, we now have a double series of nonnegative numbers, so the sum doesn't depend on the order, and so this double sum, which is $\sum |A_j| + \varepsilon$, and is at least the measure of the union of the A_j.

Let $B_1 \subset B_2 \subset \cdots \subset B_k \subset \cdots$ be an increasing sequence of measurable sets. We want to prove that

$$\left|\bigcup_{j=1}^{\infty} B_j\right| = \lim_{j\to\infty} |B_j|,$$

so we apply countable additivity to $A_k = B_k \setminus B_{k-1}$; the limit is $\sum |A_j|$, and $\bigcup B_k = \bigcup A_k$. This is a common and useful trick.

Finally, recall that $|A \cup B| = |A| + |B|$, which tells us that complements are measurable (take them inside of increasingly large cubes).

We also discussed and proved that the Borel sets (the σ-algebra generated by open sets) is a strict subset of the σ-algebra of Lesbegue-measurable sets. This is because open sets are classified by $\mathbb{Q}$, i.e. the subset of rationals in an open set characterizes it completely.[2] Then, we apply operations to these open sets that preserve this cardinality, so we get cardinality $2^{\mathbb{Q}}$ of Borel sets. But the Cantor set is Lesbegue-measurable with measure zero, and is uncountable, so all of its $2^{\mathbb{R}}$ subsets are measurable, so some of them must not be Borel.

There's also a useful condition for measurability that we'll use again when we learn about abstract measure theory.

Proposition 3.1 (Carathéodory criterion). *A set E is measurable iff for all sets A, $|A|_e = |A \cap E|_e + |A \cap E^c|_e$.*

[2] A cute way to number the rationals is $p/q \mapsto 2^p 3^q$, which creates an injection $\mathbb{Q} \hookrightarrow \mathbb{N}$.

For example, we can prove that the half-plane $H = \{x_1 > 0\}$ is measurable. This is because any covering of a set A can be split into coverings of $A \cap H$ and $A \cap H^c$ that overlap on an area of most ε. This is an alternate approach to our complicated artillery.

The idea is to prove this splitting property for open and compact sets, and therefore for G_δ sets. But we've seen that for any measurable set E, there's a G_δ set H with the same measure as E, and this can be used to prove it for all E: if E is measurable, it has the splitting property.

Conversely, if E has the splitting property, there's a measurable $H \supseteq E$ with $|H| = |E|_e$, and so H has the splitting property. But then $|H \setminus E|_e$ has measure 0, so E is measurable.

The book also covers Lipschitz transformations, i.e. those $f : \mathbb{R}^n \to \mathbb{R}^n$ such that $|f(x) - f(y)| \le C|x - y|$ for some C. For example, all linear functions are Lipschitz, and Lipschitz functions send measurable sets to measurable sets (which we'll prove later). This can be proven in the following way: first, show that it must send compact sets to compact sets, and therefore F_σ sets to F_σ sets. In particular, this means sets of measure zero are sent to sets of measure zero: the image of a cube has volume scaled by at most C^n, so if E can be covered by a countable number of cubes with total volume ε/C^n, $f(E)$ can be covered by a countable number of cubes with total volume at most ε.

So in summary, we've constructed this notion of a measure in which open and closed sets are measurable, the measurable sets are a σ-algebra, and we have countable additivity. We wanted to do this so that we could turn a space of functions into a Banach space (i.e. make sure it's a complete space), so how will we use Lesbegue measurability to make this work? We will end up using a space of integrable functions, so now we need to define integrable functions. These can be extremely discontinuous, so they will be unrelated to the topology of the basis.

We'll want a measurable $f : \mathbb{R}^n \to \mathbb{R}$ to be such that its undergraph is measurable in $\mathbb{R}^{n+1}$. Alternatively, we can approximate f from below by *simple functions*, countable weighted sums of characteristic functions. When we integrate by approximating by simple functions, we can do it in two ways: either horizontally or vertically (dividing into horizontal or vertical strips). But to do that, we need the bases to be measurable, which is only true for some functions. Thus, we will work with *Lesbegue measurable functions*.[3]

In Riemann integration, volume is signed, so integrals like $\int_0^\infty (\sin x)/x \, dx$ converge. For Lesbegue integration, we'll require things to be absolutely integrable, and therefore we're only going to discuss the Lesbegue integral of nonnegative functions. To compute the Lesbegue integral of function that are possibly negative, you need the positive and the negative sets to be measurable.

Definition. A nonnegative function $f : \mathbb{R}^n \to \mathbb{R}$ is Lesbegue measurable if for all $\alpha \in \mathbb{R}$, the set $\{x : f(x) > \alpha\}$ is measurable.

We'll want to show that these form a vector space, and that other operations such as products and certain compositions are measurable. None of this is terribly deep, but it's still useful.

First, it's equivalent to require that the sets $f < \alpha$, $f \ge \alpha$, or $f \le \alpha$ are measurable, because we can take complements, and because if $f \ge \alpha$, then $f > \alpha - 1/n$, so we're approximating a set from the outside by compact sets, and so forth. Thus, all these criteria are equivalent. You can even use a dense family of slices rather than all $\alpha \in \mathbb{R}$.

Proposition 3.2. *If f and g are measurable, then the set $S = \{x : f(x) > g(x)\}$ is measurable.*

Proof. This is because we can write

$$S = \bigcup_{q \in \mathbb{Q}} \{x : f(x) > q\} \cap \{x : q > g(x)\},$$

and the two sets on the right are measurable, and then we can take countable intersections and unions. ⊠

Corollary 3.3. *If f and g are measurable and $\lambda \in \mathbb{R}$, then $f + g$ is measurable and λf is measurable.*

This is because $f + g > \lambda$ when $f > \lambda - g$, and if $f > \alpha$, then $\lambda f > \lambda \alpha$.

Thus, measurable functions form a vector space, which is nice, because we'll want them to eventually be a Banach space.

Proposition 3.4. *If f is measurable and ϕ is continuous, then $\phi(f)$ is measurable.*

[3] In general measurable theory, we will just look at *measurable functions*.

This is because $f^{-1}(\phi^{-1}(\lambda,\infty))$ is measurable: $\phi^{-1}(\lambda,\infty)$ is open in $\mathbb{R}$ and therefore a countable union of intervals, and so the preimage of each interval under f is a measurable set (and their countable union is measurable).

So for a measurable function, the preimage of a Borel set is Lesbegue-measurable, but the preimage of a Lesbegue-measurable set might not be.

Proposition 3.5. *If f and g are measurable, then fg is measurable.*

Proof. This is a cute trick: $fg = (1/2)((f+g)^2 - (f-g)^2)$: $f+g$ is measurable and scalar multiplication preserves measurability, and if f is measurable then f^2 is just be checking the definition. ⊠

Now, we come to one of the hardest theorems of the course.

Theorem 3.6 (Egorov)**.** *Let f_k be a series of functions defined in a set E of finite measure, and suppose that f_k converges to f almost everywhere in E (i.e. the set of points where $f_k(x) \not\to f(x)$ has measure zero). Then, for any $\delta > 0$, there exists an $A \subset E$ with $|A| < \delta$ such that on $E \setminus A$, $f_k \to f$ uniformly.*[4]

Proof. We can start by replacing f_k with $f_k - f$, so that the goal is to get uniform convergence to 0. Then, since the set where $f_k \not\to f$ has measure zero, we can throw it away, and assume that $f_k(x) \to f(x)$ pointwise on all of E.

Let A_1 be the subset of E where $|f_k - f| < 1$ for $k \ge 1$, A_2 be that where $|f_k - f| < 1$ for $k \ge 2$, and so on: A_m requires this for $k \ge m$. Thus, $A_1 \subset A_2 \subset \cdots$, and the union of the A_m is E. Thus, $|A_m| \to |E|$, so $|E \setminus A_m| = |E| - |A_m| \to 0$.

We'll use $\delta/2$ of the given δ in the following way: let m_0 be the smallest m such that $|E \setminus A_m| < \delta/2$. That is, for $k > m_0$, $|f_k - f| = 1$ on A_{m_0}.

Now, let's iterate: let A_m^2 be the subset of A_{m_0} where $|f_k - f| < 1/2$ for $k > m$. Then, $A_{m+1}^2 \supset A_m$ and their union is A_{m_0}. Let m_1 be the smallest m such that $|A_{m_0} \setminus A_m^2| < \delta/4$.

Then, we repeat with the bound 1/4 and $\delta/8$, and so on and so forth, so in the limit we keep all but at most δ of the measure, and the functions uniformly converge. ⊠

Lecture 4.

Semicontinuous Functions and Review: 9/4/15

Last time, we said that an open set is characterized by its intersection with $\mathbb{Q}$. This isn't true; instead, it's characterized by its balls of rational center and rational radius. Since this is still always a countable set, then the proof that relied on this fact still goes through.

Last time, we begin discussing our space of integrable functions: we started with continuous functions, which aren't good enough, and Riemann integration isn't good enough. We want to form a set of functions that allow us to make a Banach space with the distance between two functions equal to the volume between them. We arrived at the set of measurable functions to do this.

In fact, as we'll see in the book, there are two approaches. The first considers nonnegative functions defined on some measurable set. Then, f is said to be measurable if for all λ, the set $A_\lambda = \{x : f(x) > \lambda\}$ is measurable. It's possible to take the points where $f(x) \ge \lambda$, or $<$ or $\le$, and the same set of functions is obtained.[5]

The other approach is to define f to be measurable if the set $\{(x,y) : y < f(x)\}$ is measurable in $\mathbb{R}^{n+1}$ (i.e. the undergraph of f defines a measurable set). Then, you can declare the integral of f to be the volume of that set. Alternatively, we can make a monotone approximation of a measurable function, approximating it below by step functions (e.g. first, add the characteristic function of the set where $f(x) > 1$, then $f(x) > 3/2$, then $f(x) > 11/6$, and so forth). This is reasonably intuitive, but it's not so easy to prove that this makes f measurable, since we have only a countable set of slices.

The book's approach will start with the first definition of measurability, so that when defining the integral in that way, it will have nice convergence properties as long as f is measurable; that way, we only need to worry about measurable f.

[4] $f_k \to f$ uniformly means that for an $\varepsilon > 0$, there's an N such that if $k > N$, all $f_k(x)$ are bounded by $(f(x) - \varepsilon, f(x) + \varepsilon)$.

[5] We can't just ask for $f(x) = \lambda$ to be measurable; for example, if N is a nonmeasurable set, we can take $f(x) = x$ on N and 0 outside of N, then f isn't measurable, but it passes the line test for each x.

Last lecture, we started to explore this, proving that if f and g are measurable, then λf and $f+g$ are measurable, as are fg and $\phi(f)$ for ϕ continuous.

Definition. A function f is *upper semicontinuous* at x_0 if $f(x_0) \geq \overline{\lim}_{\substack{x\to x_0\\ x\neq x_0}} f(x)$. Similarly, f is *lower semicontinuous* at x_0 if $f(x_0) \leq \underline{\lim}_{\substack{x\to x_0\\ x\neq x_0}} f(x)$.

There are various other definitions: for example, f is upper semicontinuous if whenever $x_j \to x$, then $f(x_j)$ converges to something smaller than (or equal to) $f(x_0)$.

Proposition 4.1. *If f is lower semicontinuous, then $\{x : f(x) > \lambda\}$ is open, and therefore f is measurable.*

Proof. Follows directly from the sequences definition of lower semicontinuity. ⊠

The analogous statement for upper semicontinuous functions (that $\{x : f(x) < \lambda\}$ is open, so f is measurable) is just as true.

Last time, we stated and proved the Egorov theorem: if E is measurable and has finite measure and $f_k \to f$ pointwise almost everywhere on E, then for any $\delta > 0$, one can choose out a set of measure at most δ so that on its complement in E, $f_k \to f$ uniformly. We briefly reviewed the idea of the proof in class today, but we did that last lecture, so you should look there for the proof.

The point of the uniform convergence is that the functions were chosen so that for any $\varepsilon > 0$, there's an n with $1/n < \varepsilon$, but we have a k such that $|f_\ell(x)| < 1/n$ for $\ell \geq k$, since we threw out the sets where that wasn't the case. Thus, we have uniform convergence.

A useful corollary of Egorov's theorem is Lusin's theorem: that measurable functions are almost continuous.

Theorem 4.2 (Lusin)**.** *Let f be a bounded, measurable function on a set E of finite measure. Then, given a $\delta > 0$, there's a closed set of measure $|E| - \delta$ such that $f|_K$ is continuous.*

The idea is that we'll use Egorov's theorem to approximate f by a sequence that converges uniformly on most of E, and the uniform limit must be continuous.

Since f is bounded, it can always be approximated by a sum of step functions. For each n, let $h_1, \dots, h_n$ be a partition; then, we'll let $\varphi_n = \sum_{k=1}^n \lambda_k \chi_{\{x : f(x) > h_k\}}$.

The full proof will be given next lecture.

Lecture 5.

Some More Review and the Convergence Theorems: 9/9/15

Since there's a lot of theorems and lemmas in the class, it's worth quickly reviewing the direction and ideas that this course will take.

The goal was to make the set of continuous functions (or Riemann-integrable functions) into a Banach space, which it isn't. To do that, we want to define measure. From the notion of a countable covering we arrive at outer measure; for an open set U, it's also an interior measure, as it can be well-approximated by compact subsets of U. Thus, open sets are measurable, with $|U| = |U|_e$. More generally, if a set can be approximated by open sets so that the outer measure of the difference goes to 0, then the set is called measurable. That is, if $\triangle$ denotes symmetric difference, then we want for any $\varepsilon > 0$ there to exist a B_ε such that $|A \triangle B_\varepsilon|_e < \varepsilon$. For example, open sets are measurable and complements of measurable sets are measurable.

Then, we defined Lesbegue-measurable sets, including all subsets of a set of measure zero (because the empty set is measurable). We get countable additivity for disjoint measurable sets, and if A_j is an increasing sequence of sets, then the measure of the union of the A_j is equal to the limit of their measures.

Now, we want to have an integration theory, and with $\|f\| = \int |f|$, the family of Lesbegue-integrable functions becomes a Banach space. This is constructed beginning with simple functions, which are of the form

$$\phi = \sum_{n=1}^\infty \lambda_n \chi_{E_n},$$

where each $\lambda_n \in \mathbb{R}$ and E_n is a measurable set. We know how to integrate them, i.e. in the unsurprising way, so for any function f, we can consider a sequence ϕ_k of simple functions converging to f, so that for each k, $\phi_k < f < \phi_k + 2^{-k}$ and $\phi_{k+1} \geq \phi_k$. Using the definition of measurability, you can check that measurability is equivlane to being able to be approximated this way, and that the integral of the function is the volume of the undergraph, and $\int f = \lim \int \phi_k$. This is fine as long as f is finite a.e.; and to ensure that nothing is going wrong, we just have to check that if f is finite a.e. and supported in a set E of bounded measure, then the graph of f has measure 0, and therefore it doesn't matter if we check the undergraph or the equal-and-undergraph.

So the sum of two integrable functions is integrable, and $\int af + bg = a\int f + b\int g$, because this is true for simple functions and then pass to the limit for f and g, since simple functions converge uniformly to Lesbegue-integrable functions. Moreover, the integral is additive on the sets: if E_1 and E_2 are measurable sets and f is integrable in E_1 and E_2, then f is integrable on $E_1 \cup E_2$. Once again, you prove this by taking the integrals of simple functions. Finally, if $g < f$, then simple functions again tell us that $\int g < \int f$.

There are five important theorems: Egorov and Lusin's theorems on approximation; then, three more about how limits interact with integrals. These are Fatou's lemma, the monotone convergence theorem, and the dominated convergence theorem.

Proposition 5.1. *Let f be an integrable function on E. Then, f is* absolutely continuous*: for any $\varepsilon > 0$ there exists a δ such that for any $E' \subset E$ with $|E'| < \delta$, then $\int_{E'} |f| < \varepsilon$.*

We'll approximate with simple functions: given an $\varepsilon > 0$, there's a ϕ such that $\int |f - \phi| < \varepsilon/2$, and we know it's true for ϕ with error $\varepsilon/2$, so we're good.

Definition. Let A_j be a sequence of sets. Then,

- $\liminf_{j\to\infty} A_j$ is the set of a present in all but finitely many A_j, and
- $\limsup_{j\to\infty} A_j$ is the set of a present in infinitely many A_j.

Fatou's lemma is the integral version of the following statement about sets.

Fact. If $\{A_j\}$ is a sequence of measurable sets, then $|\liminf A_j| \leq \liminf |A_j|$.

Lemma 5.2 (Fatou). *Let f_j be a sequence of nonnegative, integrable functions. Then,*

$$\int \underline{\lim}\, f_j \leq \underline{\lim} \int f_j.$$

TODO: there was a proof here, but I wasn't able to follow it. It also depended on the monotone convergence theorem, which I haven't proven yet.

The monotone convergence theorem corresponds to the following fact.

Fact. Let $\{A_j\}$ be an increasing sequence of sets. Then, $\left|\bigcup A_j\right| = \lim_{j\to\infty} |A_j|$.

The proof of the theorem with integrals follows, using the connection that the integral of a function is the volume of its undergraph.

Theorem 5.3 (Monotone convergence theorem). *Let $\{f_j\}$ be a monotone increasing sequence of nonnegative measurable functions. Then, $\int \lim f_j = \lim \int f_j$.*

Theorem 5.4 (Dominated convergence theorem). *Let $f_k \to f$ a.e. (where the f_k are integrable, but not necessarily finite) and $f_k \leq \varphi$ for some integrable φ and all k. Then, $\int f_k \to \int f$.*

The proof is straightforward: since $f = \liminf f_k$, then $\int f \leq \liminf f_k$. Then, do the same thing with $\phi - f$: $\int (\phi - f) \leq \liminf(\phi - f_k)$; putting these together, ϕ cancels out, and the result holds. In fact, since $f_k - f$ is dominated by 2ϕ (or ϕ if everything is positive), and converges a.e. to 0, so we have absolute convergence of the integral as well.

Next, we'll talk about convergence in measure, which can be a useful notion.

Definition. A sequence of functions f_k *converges in measure* to 0 if for every $\varepsilon > 0$, $|\{x : f_k(x) > \varepsilon\}| \to 0$.

Convergence pointwise a.e. implies convergence in measure, but the converse isn't true: it doesn't imply convergence almost everywhere. However, you *can* extract a subsequence which converges pointwise a.e.: for every k, choose an ℓ such that if $m \geq \ell$, $|\{x : f_m(x) > 2^{-k}\}| < 4^{-k}$; this subsequence converges to zero almost everywhere: look at $|\{f_m > 2^{-k_0}\}|$, which has measure 4^{-k_0}, so we can throw it away, and then do the same thing with k_0+1, and so forth. That means we've chosen ℓ_k such that the measure of the set where f_{ℓ_k} is greater than 2^{-k} is less than 4^{-k}. Ultimately, this means we can get our pointwise bound, except on sets as tiny as you wish.

Another thing that would be useful to prove is that the space of Lesbegue-integrable functions is complete: if for all $\varepsilon > 0$ there exists a k_0 such that $\int |f_k - f_\ell| < \varepsilon$ whenever $k, \ell > k_0$, i.e. $\{f_k\}$ is Cauchy, then there will exist a Lesbegue-integrable f such that $\int |f_n - f| \to 0$.

There's an ℓ_0 such that $\int |f_{\ell_0} - f_k| < 2^{-k_0}$ when $\ell > \ell_0$, and similarly an ℓ_1 where the bound is $2^{-(k_0+1)}$, and so on. Thus, our Cauchy sequence has been transformed into a summable series, so we look at $\int \sum |f_{\ell_{k_0}} - f_k|$. We'd like to apply Fatou, but they need to converge a.e., which is not apparent. However, we can show that they converge in measure, which will allow us to extract a subsequence whose limit will give us what we want.

Specifically, the tail of the series converges in measure. But this is easy: the sum converges, so the individual terms must get smaller and smaller. Thus, we get convergence in measure.

Here, it'll be useful to have the Chebyshev inequality, from the field of probability.

Theorem 5.5 (Chebyshev inequality)**.** *For any integrable and nonnegative f and $\lambda \geq 0$, $\int f(x) \geq \lambda \int \chi_{\{f \geq \lambda\}}$.*

Friday, we'll, uh, review and maybe discuss some more stuff.

Lecture 6.

Completeness of the Space of Integrable Functions: 9/11/15

Recall that a sequence $\{a_k\}$ is *Cauchy* if for all $\varepsilon > 0$ there exists a $k(\varepsilon)$ such that $|a_m - a_\ell| < \varepsilon$ when $\ell, m \geq k(\varepsilon)$. We can turn this into a convergent series: choose a p such that $\varepsilon < 2^{-p}$, and get the $k(p)$ for 2^{-p} and consider the subsequence $a_{k(1)}, \dots, a_{k(p)}$ and rewrite as

$$a_{k(p)} = \left(a_{k(p)} - a_{k(p-1)}\right) + \left(a_{k(p-1)} - a_{k(p-2)}\right) + \cdots + \left(a_{k(2)} - a_{k(1)}\right) + a_{k(1)}.$$

So now if $\{f_k\}$ is a Cauchy sequence of functions, we'll apply the same process to define the series

$$f_\infty = \sum \left(f_{k(p)} - f_{k(p-1)}\right) + f_{k(0)}.$$

Here, $k(p)$ is the smallest number such that

$$\int |f_m - f_\ell| < 2^{-p}$$

for all $\ell, m > k(p)$. The monotone convergence theorem tells us that this limit exists pointwise.

One can define a notion of Cauchy sequence with respect to Cauchy in measure, so-caled "Cauchy in measure." Then, it's possible to show that, as with usual convergence in measure, there's a subsequence that has a limit a.e.; then, we'll use this to show that there's a subsequence that converges almost everywhere. Finally, we'll use the dominated convergence theorem once we've produced the limit function.

Definition. A sequence $\{f_n\}$ of measurable functions is *Cauchy in measure* if for all $\varepsilon, \delta > 0$, there exists a $k_0(\varepsilon)$ such that for all $k, \ell > k_0$, $|\{x : f_k(x) - f_\ell(x) > \varepsilon\}| < \delta$.

Convergence in measure is in some sense the weakest condition, but using a series allows us to prove things, as constructed in the beginning of the lecture.

For every m, choose a $k(m)$ such that $|\{x : f_\ell(x) - f_p(x) > 2^{-k}\}| < 2^{-k}$ when $\ell, p \geq m$. Then, $\{f_{k(n)}(x)\}$ is Cauchy in m almost everywhere, and so we get a limit $f(x)$ defined pointwise a.e. In math,

$$\sum \int |f_{k(\ell)} - f_{k(m)}|$$

converges. We also may put the sum inside the integral, so this converges to $\int \phi^*$ for some ϕ^*. Moreover, we can use the Chebyshev inequality: since the series converges, we can make the tail of the series as small as we wish, and therefore given an ε, there exists a δ such that

$$\left|\{x : \sum |f_{k(\ell)} - f_{k(m)}| > \delta\}\right| \leq \varepsilon.$$

Convergence in measure is often useful for things like this. Then, apply the monotone convergence theorem, and so on.

More generally, if we want functions to also be negative, apply the argument to the positive and negative parts of the function.

When we discussed how to do Lesbegue integration, we split the undergraph into slices, and then refined our slices; this seems like a "horizontal" way to approximate the function, while Riemann sums are a "vertical" way to approximate it. So you can even do a horizontal approximation, choosing $\delta\theta$ to represent the height of a very small strip, creating a sum $\sum|f > \theta|\delta\theta$. Once again, we get something that looks somewhat like, but not entirely like, a telescoping series:

$$f(\theta_1)(\theta_2-\theta_1)+f(\theta_2)(\theta_3-\theta_2)+\cdots+f(\theta_k)(\theta_k-\theta_{k-1}).$$

In the same way as the Riemann integral, we can refine this as $k\to\infty$, and in the limit is known as the *Riemann-Stieltjes integral*; it's a pretty important concept, though not as much as the Lesbegue integral.

Next week, we'll talk about Fubini-type theorems.

Lecture 7.

Fubini's Theorem and Differentiability: 9/14/15

Today we'll discuss the topics in Chapter 6 of the textbook, including repeated integration and Fubini-type thoerems. It's a relatively straightforward chapter, and doesn't use too many deep techniques. But, as usual, everything is easier to state first with nonnegative functions, and then generalized.

Let $X\subset\mathbb{R}^n$ and $Y\subset\mathbb{R}^m$ be measurable, and let $f : X\times Y\to\mathbb{R}$ be integrable and written $f(x,y)$, with $x\in X$ and $y\in Y$. We want to know: when is it that

$$\int_{X\times Y} f(x,y)\,\mathrm{d}(x,y) \stackrel{?}{=} \int_X \mathrm{d}x\int_Y f(x,y)\,\mathrm{d}y.$$

Here, $\mathrm{d}(x,y)$ will denote the product measure.

The strategy will be to start with $f=\chi_A$, where A is measurable in $d(x,y)$, and let

$$g(x)=\int_Y \chi_A\,\mathrm{d}y = |\{y\in Y : \text{ there exists an } x\in X \text{ such that } (x,y)\in A\}|.$$

In particular, we'll discover that $\int_Y g(y)\,\mathrm{d}y = \int_{X\times Y}\chi_A\,\mathrm{d}(x,y)$. We'll call $g(y)$ the *trace* of $f(x,y)$ in the y-direction.

We'll prove this by starting with rectangles, and then successively generalizing.

- First, to finite unions of nonoverlapping rectangles, which preserves compactness.
- Then, to characteristic functions open sets in $\mathbb{R}^n\times\mathbb{R}^m$, because each of these can be covered by an increasing sequence of finite disjoint rectangles. This uses the monotone convergence theorem, and the fact that
$$\int_X \mathrm{d}x\int_Y \chi_{A_k}\,\mathrm{d}y$$
is a piecewise constant function, and therefore measurable, so applying the monotone convergence theorem in x handles it also.
- Then, it can be applied to sets of measure zero, using the monotone convergence theorem on functions which decrease monotonically on G_δ sets, and more or less the same argument goes through. This requires showing that if $E\subset\mathbb{R}^{n+m}$ has measure zero, then almost every trace has measure zero (and then the integral has measure zero). This can be proven by covering a set A of measure zero with a decreasing sequence of open sets $\{U_k\}$ converging to A, so $\chi_{U_k}\to\chi_A$ and therefore their integrals also converge; therefore the integral on any trace of χ_A must be zero almost everywhere (since this is true for each χ_{U_k} as well).

This isn't the most exciting stuff, I suppose; it's a pretty straightforward theorem.

So every measurable set can be written as a G_δ set along with a set of measure zero, and this along with the above sequence of generalizations shows that if $A\subset\mathbb{R}^{m+n}$ is measurable, then almost every trace is measurable with respect to dy, and the measure of the trace at $x \in \mathbb{R}^n$ is

$$g(x) = \int \chi_A(x,y)\,\mathrm{d}y,$$

which defines a measurable function g. Moreover,

$$\int g(x)\,\mathrm{d}x = \int \mathrm{d}x \int \chi_A(x,y)\,\mathrm{d}y = \iint\limits_A \chi_A(x,y)\,\mathrm{d}(x,y). \tag{9.14.1}$$

But this is immediately true for finite linear combinations $\sum \lambda_j \chi_{A_j}$, and any measurable f can be well approximated by these simple functions s_n, and in fact can be approximated monotonically. Therefore the integrals are also monotonic, and thus converge, so (9.14.1) applies to all measurable f.

Moreover, one doesn't have to assume f is integrable with respect to every measure; instead, you can assume the integral on either side is finite, and then show that the integrals are the same. But you do want to know that f is measurable in x and in y; not only is this necessary, but once this is established, the finiteness of any one of the three integrals in (9.14.1) being finite implies the other two are, and that they're equal.

The notion that the integral of a function is the area under its graph leads to the theorem that a function whose undergraph is measurable is itself an integrable function. This ties back to Fubini's theorem because we're measusing in one more variable, so let $x \in \mathbb{R}^n$ and $y \in \mathbb{R}$. Then, we consider the set $\{f(x) > y\}$, and we want to show that it's measurable with respect to dx.

That's, uh, most of the chapter. There's also Tonelli's theorem, but that's sort of already been covered.

There's an important application of Fubini's theorem to convolutions. Suppose $f, g : \mathbb{R}^n \to \mathbb{R}$ are measurable, positive functions and have finite integral, and we want to look at the convolution

$$f * g = \int f(y) g(x-y)\,\mathrm{d}y.$$

The idea is to approximate $g(y)$ by nicer functions. First, replace $g(x)$ with $1/\varepsilon^n g(x/\varepsilon)$, to ensure that the integral stays equal to 1 even when we squeeze the graph, in a sense making g pointier and pointier.

For example, if we take $g = \chi_{B_1(0)}/|B_1(0)|$ (so that g has volume 1); then, g_ε is the characteristic function of $B_\varepsilon(0)$, normalized to have integral 1.

Thus, when we look at its convolution with f at a point x,

$$(f * g_\varepsilon)(x) = \int f(x-y) g_\varepsilon(x)\,\mathrm{d}y$$

is the average of f on $B_\varepsilon(x)$! Well, that's imprecise, but since $\int_{B_\varepsilon(0)} g = 1$, then it's as reasonable a definition as we'll get. Then, Fubini's theorem tells us that

$$\iint f(x-y) g_\varepsilon(y)\,\mathrm{d}y\,\mathrm{d}x$$

can be integrated in any order we want; if you integrate with respect to y first, it becomes $\int g_\varepsilon(y)\,\mathrm{d}y \int f(x)\,\mathrm{d}x$, and therefore the integral of the average is the same![6]

The next chapter is one of the professor's favorites, unlike Fubini's theorem: differentiability. You remember that if one defines

$$F(t) = \int_a^t f(x)\,\mathrm{d}x,$$

then $F'(t) = f(t)$ for well-behaved f. In a more discrete way, if f is well-behaved and $\varepsilon \to 0$,

$$\frac{1}{2\varepsilon} \int_{t-\varepsilon}^{t+\varepsilon} f(x)\,\mathrm{d}x \longrightarrow f(t).$$

But more generally, in $\mathbb{R}^n$ with $n \geq 1$, we would rewrite this as

$$\frac{1}{|B_\varepsilon(t)|} \int\limits_{B_\varepsilon(t)} f(x)\,\mathrm{d}x \longrightarrow f(t), \tag{9.14.2}$$

[6] If your function isn't nonnegative, then you have to add absolute values, and then the average decreases the L^1-norm, rather than preserving it.

which holds when f is continuous. Notice the connection with convolutions!

Anyways, we know these are true for continuous functions. When we generalize to integrable functions, what changes? When is it still true almost everywhere?

Let A be a measurable function, so we can first look at (9.14.2) for χ_A. Then, $\int_{B_\varepsilon(t)} \chi_A(x)\,\mathrm{d}x \to \chi_A(t)$ would imply $|A \cap B_\varepsilon|/|B_\varepsilon| \to 1$ as $\varepsilon \to 0$ a.e. Measurable sets can be quite poorly behaved in general, though. This is caled the theory of *Lesbegue differentiation*.

In general, the average of f over $B_\varepsilon(t)$ is denoted

$$\fint_{B_\varepsilon(t)} f(x)\,\mathrm{d}x = \frac{1}{|B_\varepsilon(t)|} \int_{B_\varepsilon(t)} f(x)\,\mathrm{d}x.$$

We know we can approximate integrable functions with continuous ones, so the natural thing to do is to approximate the average around x with a continuous function in x. One can invoke a powerful analytic tool of Littlewood and Hardy.

Definition. Let f be a nonnegative integrable function, and let

$$Mf(y) = \sup_{\lambda > 0} \fint_{B_\lambda(y)} f(x)\,\mathrm{d}x.$$

This is called the *Littlewood-Hardy maximal function* of f.

Then, we'll have the following results.

(1) If $f \in L^\infty$ (i.e. f is bounded), then $Mf \in L^\infty$ as well, and $\|Mf\|_{L^\infty} \le \|f\|_{L^\infty}$ (since the supremum will be averaged with things at most as large as it).

(2) Mf might not be L^1, even if f is, e.g. $f = \chi_{B_r(0)}$. If it is L^1, though, then we have a Chebyshev inequality

$$|\{x : Mf(x) > \lambda\}| \le \frac{1}{\lambda} \int |f|.$$

(3) If $f, Mf \in L^1$ and $p1$, then $\|Mf\|_{L^p} \le C\|f\|_{L^p}$, for some C independent of f.

Now we will get down to the business of proving things.

Lemma 7.1 (Rising Sun)**.** *Let f be a positive measurable function and*

$$A_\lambda = \left\{ \sup \frac{1}{h} \int_x^{x+h} f(t)\,\mathrm{d}t > \lambda \right\}.$$

Then, $|A_\lambda| \le 1/\lambda$.