**Lecture 1 - 08/25/2014**

Book Recommendations

- For grad students, use Folland. In-depth, comprehensive, challenging.
- Also good but simpler is Cohn. Cohn is a very nice read.
- Rudin is good, but very dense.

**Quotes** - "My favourite book is Folland." "These books are all isomorphic."

Here's some motivation for the development of the Lebesgue measure/Lebegue integral:
Challenge - Let $(f_n)$ be a sequence of functions defined on $[0,1]$, $0 \leqslant f_n \leqslant 1$ for all $n$. Show that if $f_n \to 0$ pointwise, then $\lim\limits_{n\to\infty} f_n \int_0^1 f_n = 0$.

Using Riemann integration, this is easy if $f_n \to 0$ uniformly. Not true if we only have pointwise convergence. However, the problem becomes easy again if we use the theory of Lebesgue integration.

Intuition - The Riemann integral is calculated by subdividing the domain of a function, then drawing rectangles to approximate the area under a curve. The Lebesgue integral is calculated by subdividing the *range* of a function. Approximate $f$ by a function of the form $\sum\limits_{i=1}^{n} a_i \chi_{A_i}$, where $\chi_S(x) = \begin{cases} 1 & x \in S \\ 0 & x \notin S \end{cases}$, where the sets $A_i$ are disjoint and "nice" and $f \approx a_i$ on each $A_i$. We then approximate

$$\int_a^b f \approx \sum_{i=1}^{n} a_i \cdot \text{"length}(A_i)\text{"}$$

For example, we might have $a_1 < \cdots < a_n$, $A_i = f^{-1}([a_i, a_{i+1}))$.

To get some intuition for this analogy, consider a banker counting money. One way to do this is to add the bills together in sequence. This is like Riemann integration. Lebesgue integration is analogous to sorting the bills into stacks of different denomination, then counting the size of each stack.

We will develop a notion of "size" for a set $A$, denoted by $\lambda(A)$. In one dimension, this will correspond to length. In two dimensions, this will correspond to area. In three dimensions, this will correspond to volume, etc.

There are some things we must be careful about, lest we get in trouble.

Banach-Tarski Paradox - There is a positive integer $n$, disjoint sets $A_1, \ldots, A_n \subseteq \mathbb{R}^3$, and maps $R_1, \ldots, R_n : \mathbb{R}^3 \to \mathbb{R}^3$ that are combinations of rotations, reflections, and translations such that $\bigcup\limits_{1\leqslant i\leqslant n} A_i = B(0,1)$ and

$$\bigcup_{1\leqslant i\leqslant n} R_i(A_i) = B(0,1) \cup B(2,1)$$

**Quote** - "With this problem, I've solved world hunger. Give me one loaf of bread, I'll give you two!"

Issue - Although it seems like we're doubling the volume of the ball, there ends up being no sensible notion of volume for the sets $A_i$, so we can't identify the volume of the ball with the sum of the volumes of the pieces. We'll formalize this later.
Remark - This can be done with $n = 5$ (and this is optimal).

Definition - Let $X$ be a set. We say $\Sigma \subseteq \mathcal{P}(X)$ is a $\sigma$-algebra on the set $X$ if

1. $\varnothing \in \Sigma$
2. $A \in \Sigma \implies A^c \in \Sigma$
3. $(A_n) \subseteq \Sigma \implies \bigcup_n A_n \in \Sigma$

Intuition - Think of $\Sigma$ as all subsets of $X$ to which we can assign a meaningful notion of volume. These will be our "measurable" sets.
Remark - Conditions 1 and 2 imply that $X \in \Sigma$. Conditions 2 and 3 imply that $\Sigma$ is closed under countable intersections.

Example - $\Sigma = \{\varnothing, X\}$, $\Sigma = \mathcal{P}(X)$.

Proposition - If $\Sigma_1$ and $\Sigma_2$ are $\sigma$-algebras on $X$, then so is $\Sigma_1 \cap \Sigma_2$. In fact, if $\{\Sigma_\alpha\}_{\alpha \in A}$ is a collection of $\sigma$-algebras on a set $X$, then so is $\bigcap_{\alpha \in A} \Sigma_\alpha$.
Proof - Easy exercise. □

Definition - Let $\mathcal{E} \subseteq \mathcal{P}(X)$ be arbitrary. Denote by $\sigma(E)$ the $\sigma$-algebra on $X$ *generated by* $\mathcal{E}$. This is defined to be the intersection of all $\sigma$-algebras on $X$ that contain $\mathcal{E}$.

By the previous proposition, this definition is justified. Intuitively, $\sigma(\mathcal{E})$ is the smallest $\sigma$-algebra on $X$ that contains $\mathcal{E}$.

Example/Definition - Take $X = \mathbb{R}^d$, $\mathcal{E} = \{\text{open sets}\}$. We call $\sigma(\mathcal{E})$ the *Borel* $\sigma$-algebra on $X$ and denote it by $\mathcal{B}(\mathbb{R}^d)$.
Remark - $\mathcal{B}(\mathbb{R}^d) = \sigma(\{\text{hypercubes in } \mathbb{R}^d\})$ (exercise).
Remark - $\mathcal{B}(\mathbb{R}^d) = \sigma(\{\text{closed sets}\})$.

Definition - Let $X$ be a set, $\Sigma$ a $\sigma$-algebra on $X$. We say $\mu : \Sigma \to [0, \infty]$ is a *positive measure* on $(X, \Sigma)$ if

1. $\mu(\varnothing) = 0$.
2. Given disjoint sets $(A_n) \subseteq \Sigma$, we have $\mu(\bigcup_{n \geqslant 1} A_n) = \sum_{n \geqslant 1} \mu(A_n)$

Remark - In general, (2) doesn't imply (1). Why? Given $A \in \Sigma$, we have

$$\mu(A) = \mu(A \cup \varnothing) = \mu(A) + \mu(\varnothing)$$

We'd like to conclude that $\mu(\varnothing) = 0$, but this is only true if $\mu(A) < \infty$. Thus, 2 implies 1 if there is $A \in \Sigma$ such that $\mu(A) < \infty$.

Example (Counting measure) - Let $X$ be any set, $\Sigma = \mathcal{P}(X)$, and define $\mu(A) = |A|$, the cardinality of $A$. That this is a measure is left as an exercise.
Example ($\delta$ measure) - Let $a_0 \in X$, $\Sigma = \mathcal{P}(X)$. Define

$$\delta_{a_0}(X) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

That this is a measure is left as an exercise. We call this the $\delta$ measure at $a_0$.

**Lecture 2 - 08/27/2014**

Goal - Construct the Lebesgue measure, which will be a measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. This measure will satisfy

$$\lambda((a_1, b_1) \times \cdots \times (a_d, b_d)) = \prod_{1 \leqslant i \leqslant d} (b_i - a_i)$$

for all $a_i < b_i$. That is, the measure will agree on cells in $\mathbb{R}^d$.

Goal 2 - Define a more robust notion of integration. Abstractly, given a measure space $(X, \Sigma, \mu)$:

1. If $s : X \to \mathbb{R}$ is of the form $s = \sum_{1 \leqslant i \leqslant N} a_i \chi_{A_i}$, where the sets $A_i \in \Sigma$ are disjoint, define $\int_X s \, d\mu = \sum_{1 \leqslant i \leqslant N} a_i \mu(A_i)$.
2. Given an arbitrary function $f : X \to \mathbb{R}$, approximate $f$ by functions from (1) to define $\int_X f \, d\mu$.

**Construction of Lebesgue Measure**

Definition - We say $I \subseteq \mathbb{R}$ is a *cell* if it is a finite interval, i.e. $I$ is of the form $(a, b)$, $(a, b]$, $[a, b)$, or $[a, b]$, where $a, b \in \mathbb{R}$.
Definition - We say $I \subseteq \mathbb{R}^d$ is a *cell* if $I = I_1 \times \cdots \times I_d$, where each $I_i$ is a cell in $\mathbb{R}$. Given a nonempty cell $I$ in $\mathbb{R}^d$, define

$$\ell(I) = \prod_{1 \leqslant i \leqslant d} (\sup(I_i) - \inf(I_i))$$

and define $\ell(\varnothing) = 0$. (Intuitively, this is volume.)

Definition - Given $A \subseteq \mathbb{R}^d$ arbitrary, define

$$\lambda^*(A) = \inf\{\sum_{n \geqslant 1} \ell(I_n) \mid A \subseteq \bigcup_{n \geqslant 1} I_n, \; I_n \text{ cells}\}$$

We call $\lambda^*$ the *Lebesgue outer measure*. Although $\lambda^*$ is defined on all of $\mathcal{P}(\mathbb{R}^d)$, it is only countably additive on a proper subset of $\mathcal{P}(\mathbb{R}^d)$ (as we shall see).

Properties of $\lambda^*$

1. $\lambda^*(\varnothing) = 0$
2. $A \subseteq B \implies \lambda^*(A) \leqslant \lambda^*(B)$ (monotonicity)
3. If $A_i \subseteq \mathbb{R}^d$, then $\lambda^*(\bigcup_{i \geqslant 1} A_i) \leqslant \sum_{i \geqslant 1} \lambda^*(A_i)$ (countable subadditivity)

Proof - (1) and (2) are immediate. (**Quote** - "For the entire semester, there's only one trick we keep using, which we'll use now.") For (3), let $\epsilon > 0$. For each $i$, let $(I_{i,n})$ be a cover of $A_i$ by cells such that

$$\sum_{n \geqslant 1} \ell(I_{i,n}) \leqslant \lambda^*(A_i) + \tfrac{\epsilon}{2^i}$$

Clearly $\{I_{i,n}\}_{i,n \in \mathbb{N}}$ is a countable cover of $\bigcup_{i \geqslant 1} A_i$ by cells, so by definition, we have

$$\lambda^*(\bigcup_{i \geqslant 1} A_i) \leqslant \sum_{i,n} \ell(I_{i,n}) \leqslant \sum_i (\lambda^*(A_i) + \tfrac{\epsilon}{2^i}) = \sum_{i \geqslant 1} \lambda^*(A_i) + \epsilon$$

Since $\epsilon > 0$ was arbitrary, let $\epsilon \to 0^+$ to get $\lambda^*(\bigcup_{i \geqslant 1} A_i) \leqslant \sum_{i \geqslant 1} \lambda^*(A_i)$. □

Next time, we'll show two things. First, that $\lambda^*(I) = \ell(I)$ for any cell $I$. Next, we'll show that $\lambda^*$ has a property called *separated additivity*. That is, if the distance between two sets $A, B \subseteq \mathbb{R}^d$ is positive, then $\lambda^*(A \cup B) = \lambda^*(A) + \lambda^*(B)$.

**Lecture 3 - 08/29/2014**

Definition - We say $A, B \subseteq \mathbb{R}^d$ are *separated* if $d(A,B) > 0$, where

$$d(A,B) = \inf\{|a-b| \mid a \in A,\ b \in B\}$$

Proposition 1 - If $A, B \subseteq \mathbb{R}^d$ with $d(A,B) > 0$, then $\lambda^*(A \cup B) = \lambda^*(A) + \lambda^*(B)$.
Proof - We may assume $\lambda^*(A) + \lambda^*(B) < \infty$ (otherwise, the result is clear by monotonicity). By subadditivity, we have $\lambda^*(A \cup B) \leqslant \lambda^*(A) + \lambda^*(B)$, so it suffices to show that $\lambda^*(A \cup B) \geqslant \lambda^*(A) + \lambda^*(B)$.

Denote $\delta = d(A,B)$ and let $\{I_k\}$ be a cover of $A \cup B$ by cells. By subdividing each $I_k$ into cells of diameter at most $\frac{\delta}{2}$, we obtain a new collection $\{J_k\}$ of cells. Note that $\sum\limits_k \ell(I_k) \geqslant \sum\limits_k \ell(J_k)$ since the subdivision process may result in new cells that are duplicated (and we only count each cell once).

Since $\operatorname{diam}(J_k) < \frac{\delta}{2}$ for each $k$, then $J_k \cap A \neq \varnothing$ implies $J_k \cap B = \varnothing$ and vice-versa. Now partition $\{J_k\} = \{J_k'\} \cup \{J_k''\}$, where $J_k' \cap A \neq \varnothing$ and $J_k'' \cap A = \varnothing$ for each $k$. Now

$$\bigcup_k J_k' \supseteq A, \quad \bigcup_k J_k'' \supseteq B \implies \sum_k \ell(I_k) \geqslant \sum_k \ell(J_k) = \sum_k \ell(J_k') + \sum_k \ell(J_k'') \geqslant \lambda^*(A) + \lambda^*(B)$$

Taking the infimum over all covers $\{I_k\}$, we have $\lambda * (A \cup B) \geqslant \lambda^*(A) + \lambda^*(B)$, as desired. □

Remark - For sets $A$ and $B$ that aren't separated, we don't have disjoint additivity. Intuitively, the problem is with the boundary of $A$ and $B$. If $A$, $B$ are sufficiently complicated, then there is no way to divide a cover of $A \cup B$ into a cover of $A$ and a cover $B$ (like we did above). Covering $A$ may necessarily entail covering some of $B$ and/or vice-versa, so we get double-counting.

Proposition 2 - If $I$ is a cell, then $\lambda^*(I) = \ell(I)$.
Lemma - $\lambda^*(A) = \inf\{\sum\limits_n \ell(I_n) \mid A \subseteq \bigcup\limits_n I_n,\ I_n \text{ open cells}\} =: \mu^*(A)$.
Proof - Clearly $\lambda^*(A) \leqslant \mu^*(A)$. We need to show that $\mu^*(A) \leqslant \lambda^*(A)$. We may assume $\lambda^*(A) < \infty$, otherwise this is clear.

Let $\epsilon > 0$ and let $(I_n)$ be a cover of $A$ by cells such that $\sum\limits_n \ell(I_n) \leqslant \lambda^*(A) + \epsilon$. For each $n$, let $J_n \supseteq I_n$ be an open cell with $\ell(J_n) - \ell(I_n) < \frac{\epsilon}{2^n}$. Now

$$\mu^*(A) \leqslant \sum_n \ell(J_n) \leqslant \sum_n \ell(I_n) + \epsilon \leqslant \lambda^*(A) + 2\epsilon$$

Since $\epsilon > 0$ was arbitrary, then $\mu^*(A) \leqslant \lambda^*(A)$, as desired. □

Proof (of Proposition 2) - First assume that $I$ is closed. Clearly $\lambda^*(I) \leqslant \ell(I)$ by definition. We need to show $\lambda^*(I) \geqslant \ell(I)$. By Lemma 1, it's enough to consider covers of $I$ by open cells.

Say $(I_n)$ is a cover of $I$ by open cells in $\mathbb{R}^d$. We need to show that $\sum\limits_{n \geqslant 1} \ell(I_n) \geqslant \ell(I)$. Since $I$ is closed, it is compact, so there is $N$ such that $I \subseteq \bigcup\limits_{1 \leqslant n \leqslant N} I_n$. Extend the faces of the cells $\{I_n\}$ to hyperplanes and use these to subdivide $I$ into new cells $(J_n)$. Clearly $\sum\limits_n \ell(J_n) = \ell(I)$ (by the distributive law for multiplication in $\mathbb{R}$), and $\sum\limits_n \ell(J_n) \leqslant \sum\limits_n \ell(I_n)$. Thus, $\sum\limits_n \ell(I_n) \geqslant \ell(I)$. Taking the infimum over covers of $I$ by open cells, we have $\lambda^*(I) \geqslant \ell(I)$, as desired.

If $I$ is not closed, let $J \subseteq I$ be a closed cell such that $\ell(J) \geqslant \ell(I) - \epsilon$. Then

$$\lambda^*(I) \geqslant \lambda^*(J) = \ell(J) \geqslant \ell(I) - \epsilon$$

Taking $\epsilon \to 0^+$, we have the desired result. □

Remark (translation invariance) - For any cell $I$ and point $x_0$ in $\mathbb{R}^d$, $\ell(I) = \ell(I + x_0)$. Therefore, $\lambda^*(A) = \lambda^*(A + x_0)$ for any set $A$ and point $x_0$. That is, the Lebesgue outer measure is *translation invariant*.
Remark - There's another way to obtain $\lambda^*$. Instead of using the volume of cubes, try and define a notion of volume by requiring it to be translation invariant (among other things). Doing this, we recover the Lebesgue measure. More generally, this can be applied to topological groups to define something called *Haar measure*.

We return to the construction of the Lebesgue measure, introducing Carathéodory's criterion and abstracting our discussion.

Definition - We say $\mu^*$ is an *outer measure* on $X$ if $\mu^* : \mathcal{P}(X) \to [0, \infty]$ and

1. $\mu^*(\varnothing) = 0$
2. $A \subseteq B \implies \mu^*(A) \leqslant \mu^*(B)$
3. $\mu^*(\bigcup_n A_n) \leqslant \sum_n \mu^*(A_n)$

Example - $\lambda^*$ is an outer measure.

Theorem (Carathéodory) - Let $\mu^*$ be an outer measure on $X$. Define

$$\Sigma = \{E \mid \text{for all } A \subseteq X,\ \mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)\}$$

Then

1. $\Sigma$ is a $\sigma$-algebra.
2. $\mu \restriction_\Sigma$ is a measure.

Proof - Next time.

**Lecture 4 - 09/01/2014**

Theorem (Carathéodory) - Let $\mu^*$ be an outer measure on $X$. Define

$$\Sigma = \{E \mid \text{for all } A \subseteq X,\ \mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)\}$$

Then

1. $\Sigma$ is a $\sigma$-algebra, the ($\mu^*$-)*measurable* subsets of $X$.
2. $\mu \restriction_\Sigma$ is a measure.

Proof - We first show that $\varnothing \in \Sigma$ and $\Sigma$ is closed under complementation.

- $\varnothing \in \Sigma$: For any $A \subseteq X$, we have $\mu^*(A) = 0 + \mu^*(A) = \mu^*(A \cap \varnothing) + \mu^*(A \cap X)$.
- Given $E \in \Sigma$, we have $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c) = \mu^*(A \cap E^c) + \mu^*(A \cap (E^c)^c)$.

To show that $\Sigma$ is closed under countable unions, we first prove some facts about finite unions and $\mu^*$.

**Claim 1** $E, F \in \Sigma \implies E \cup F \in \Sigma$. (Thus, $\Sigma$ is closed under finite unions.)

**Proof** Let $A \subseteq X$. Then, appealing to measurability of $E, F$, we have

$$\begin{aligned}
&\mu^*(A \cap (E \cap F)) + \mu^*(A \cap (E \cap F)^c) \\
&= \mu^*(A \cap (E \cup F) \cap E) + \mu^*(A \cap (E \cup F) \cap E^c) + \mu^*(A \cap (E \cap F)^c) \\
&= \mu^*(A \cap E) + \mu^*(A \cap F \cap E^c) + \mu^*(A \cap F^c \cap E^c) \\
&= \mu^*(A \cap E) + \mu^*(A \cap E^c) \\
&= \mu^*(A)
\end{aligned}$$

□

**Claim 2** If $E_1, \ldots, E_n \in \Sigma$ are pairwise-disjoint and $A \subseteq X$, then

$$\mu^*(A \cap (\bigcup_{1 \leqslant i \leqslant N} E_i)) = \sum_{1 \leqslant i \leqslant N} \mu^*(A \cap E_i)$$

**Proof** By induction, it's enough to consider $N = 2$. Let $E, F \in \Sigma$ be disjoint, $A \subseteq X$. Now

$$\mu^*(A \cap (E \cup F)) = \mu^*(A \cap (E \cup F) \cap E) + \mu^*(A \cap (E \cup F) \cap E^c) = \mu^*(A \cap E) + \mu^*(A \cap F)$$

□

To show that $\Sigma$ is closed under countable unions (proving (1)), it's enough by Claim 1 to show that $\Sigma$ is closed under finite disjoint unions. Let $(E_n) \subseteq \Sigma$ pairwise-disjoint, $A \subseteq X$. We want $\bigcup_n E_n \in \Sigma$.

By subadditivity, it's enough to show that $\mu^*(A) \geqslant \mu^*(A \cap \bigcup_n E_n) + \mu^*(A \cap (\bigcup_n E_n)^c)$. By Claim 2 and monotonicity, we have

$$\begin{aligned}
\mu^*(A) &= \mu^*(A \cap \bigcup_{1 \leqslant n \leqslant N} E_n) + \mu^*(A \cap (\bigcup_{1 \leqslant n \leqslant N} E_n)^c) \\
&\geqslant \mu^*(A \cap \bigcup_{1 \leqslant n \leqslant N} E_n) + \mu^*(A \cap (\bigcup_{n \geqslant 1} E_n)^c) \\
&= \sum_{1 \leqslant n \leqslant N} \mu^*(A \cap E_n) + \mu^*(A \cap (\bigcup_{n \geqslant 1} E_n)^c)
\end{aligned}$$

Taking the limit $N \to \infty$, we have (by subadditivity)

$$\mu^*(A) \geqslant \sum_{n\geqslant 1} \mu^*(A \cap E_n) + \mu^*(A \cap (\bigcup_{n\geqslant 1} E_n)^c) \geqslant \mu^*(A \cap (\bigcup_{n\geqslant 1} E_n)) + \mu^*(A \cap (\bigcup_{n\geqslant 1} E_n)^c)$$

Thus, we have equality above, which proves the result. Moreover, an immediate consequence of this proof is that $\mu^*$ is countably additive on $\Sigma$ (take $A = \bigcup_{n\geqslant 1} E_n$). $\square$

Let's return to the construction of the Lebesgue measure on $\mathbb{R}^d$. Let $\lambda^*$ be the Lebesgue outer measure. By Carathéodory,

$$\mathcal{L} := \mathcal{L}(\mathbb{R}^d) = \{E \subseteq \mathbb{R}^d \mid \text{for all } A \subseteq \mathbb{R}^d,\ \lambda^*(A) = \lambda^*(A \cap E) + \lambda^*(A \cap E^c)\}$$

is a $\sigma$-algebra and $\lambda := \lambda^* \restriction_\Sigma$ is a measure, the *Lebesgue measure*.

This seems promising, but there's one more problem. We don't know what $\mathcal{L}$ is! It may be trivial!

Proposition - Let $I$ be a cell. Then $I \in \mathcal{L}$.
Corollary - $\mathcal{B}(\mathbb{R}^d) \subseteq \mathcal{L}(\mathbb{R})$.
Remark - $\mathcal{B}(\mathbb{R}^d) \subsetneq \mathcal{L}(\mathbb{R}^d) \subsetneq \mathcal{P}(\mathbb{R}^d)$. In fact, $|\mathcal{B}(\mathbb{R}^d)| = |\mathbb{R}|$ and $|\mathcal{L}(\mathbb{R}^d)| = |\mathcal{P}(\mathbb{R}^d)|$.

Proof (of Proposition) - We need to show that for all $A$, $\lambda^*(A) \geqslant \lambda^*(A \cap I) + \lambda^*(A \cap I^c)$. Let $\epsilon > 0$. Choose a cell $I_1 \subseteq I$ such that $\lambda^*(I \backslash I_1) < \epsilon$ and $d(I_1, I^c) > 0$. By separated additivity and monotonicity of $\lambda^*$, we have

$$\lambda^*(A \cap I_1) + \lambda^*(A \cap I^c) = \lambda^*(A \cap (I_1 \cup I^c)) \leqslant \lambda^*(A) \qquad (*)$$

Moreover,

$$\lambda^*(A \cap I) \leqslant \lambda^*(A \cap I_1) + \lambda^*(A \cap (I \backslash I_1)) \leqslant \lambda^*(A \cap I_1) + \epsilon$$

Rearranging, we have $\lambda^*(A \cap I_1) \geqslant \lambda^*(A \cap I) - \epsilon$. By $(*)$, we have

$$\lambda^*(A) \geqslant \lambda^*(A \cap I_1) + \lambda^*(A \cap I^c) \geqslant \lambda^*(A \cap I) + \lambda^*(A \cap I^c) - \epsilon$$

Let $\epsilon \to 0^+$ to get $\lambda^*(A) \geqslant \lambda^*(A \cap I) + \lambda^*(A \cap I^c)$. $\square$

We've now shown that the Lebesgue measure exists, is nontrivial, and agrees with our intuitive notion of volume for cells. Next time, we'll show that the Lebesgue measure is the unique measure that satisfies our requirements.

**Lecture 5 - 09/05/2014**

Theorem (Uniqueness) - If $\mu$ is a measure on $(\mathbb{R}^d, \mathcal{L})$ (or $(\mathbb{R}^d, \mathcal{B})$) and $\mu(I) = \lambda(I)$ for all cells $I$, then

$$\mu(A) = \lambda(A)$$

for all $A \in \mathcal{L}$ (or $\mathcal{B}$).
Proof - =We proceed in steps.

1. Let $A \in \mathcal{L}$ be arbitrary. Let $(I_n)$ be a cover of $A$ by cells. Then by countable subadditivity and monotonicty,
$$\mu(A) \leqslant \sum_n \mu(I_n) = \sum_n \lambda(I_n) = \sum_n \ell(I_n)$$
Taking the infimum over all possible covers, we have $\mu(A) \leqslant \lambda(A)$.

2. Now suppose that $A \in \mathcal{L}$ is bounded, say $A \subseteq I$, a bounded cell. By the previous step,
$$\mu(I \backslash A) \leqslant \lambda(I \backslash A) \implies \mu(I) - \mu(A) \leqslant \lambda(I) - \lambda(A) \implies \lambda(A) \leqslant \mu(A)$$
By the previous step, we have $\mu(A) = \lambda(A)$.

3. For arbitrary $A \in \mathcal{L}$, write $A = \bigcup_{n \geqslant 1} (A \cap (-n,n)^d)$. Now
$$\mu(A) = \lim_{n\to\infty} \mu(A \cap (-n,n)^d) = \lim_{n\to\infty} \lambda(A \cap (-n,n)^d) = \lambda(A),$$
which finishes the proof. □

Our approach above used a neat trick, but now we'll consider a more general, abstract approach to uniqueness.

**Uniqueness Abstractly**

Let $X$ be some set, $\Sigma$ a $\sigma$-algebra on $X$, $\mathcal{C} \subseteq \mathcal{P}(X)$. Let $\mu, \nu$ be measures on $(X, \Sigma)$ such that $\mu(X) = \nu(X) < \infty$. Suppose that $\mu = \nu$ on $\mathcal{C}$. Does this imply that $\mu = \nu$ on $\sigma(\mathcal{C})$.

As it turns out, the answer is no! For example, consider two sets $A$ and $B$. We might have $\mu(A \backslash A \cap B) = \mu(B \backslash A \cap B) = 0$ and $\mu(A \cap B) = 1$, while $\nu(A \backslash A \cap B) = \nu(B \backslash A \cap B) = 1$, while $\nu(A \cap B) = 0$. Then $\mu = \nu$ on $\{A, B\}$, but $\mu$ and $\nu$ disagree on $A \cap B$.

To prove that $\mu = \nu$ on $\sigma(\mathcal{C})$, we need $\mathcal{C}$ to be closed under intersections! Now denote $\mathcal{D} = \{A \mid \mu(A) = \nu(A)\}$. Does $\mathcal{D}$ have the properties of a $\sigma$-algebra that we want?

1. $\varnothing \in \mathcal{D}$ is clear.

2. $A \in \mathcal{D} \implies A^c \in \mathcal{D}$ follows from $\mu(X) = \nu(X) < \infty$.

3. We can't generally show that $\mathcal{D}$ is closed under finite unions (let alone countable) unless we know that $\mathcal{D}$ is closed under finite intersections! Instead, let's consider a property that we know holds of $\mathcal{D}$.

3′ If $(A_n) \subseteq \mathcal{D}$ is *increasing*, then $\bigcup_n A_n \in \mathcal{D}$.

2′ While we're at it, let's generalize (2): If $A, B \in \mathcal{D}$ with $A \subseteq B$, then $B \backslash A \in \mathcal{D}$. So, not only is $\mathcal{D}$ closed under complements, but it's closed under *relative* complements.

The discussion above motivates a number of definitions.

Definition - We say $\Pi \subseteq \mathcal{P}(X)$ is a $\pi$-system $\Pi \neq \varnothing$ and

$$A, B \in \Pi \implies A \cap B \in \Pi$$

Definition - We say $\Lambda \subseteq \mathcal{P}(X)$ is a $\lambda$-system if

1. $X \in \Lambda$
2. $A, B \in \Lambda,\ A \subseteq B \implies B \backslash A \in \Lambda$
3. $(A_n) \subseteq \Lambda,\ A_n \subseteq A_{n+1} \implies \bigcup_n A_n \in \Lambda.$

Theorem (Dynkin) - If $\Pi$ is a $\pi$-system and $\Lambda$ is a $\lambda$-system, then

$$\Lambda \supseteq \Pi \implies \Lambda \supseteq \sigma(\Pi)$$

Corollary - If $\mu, \nu$ are finite measures that agree on $\Pi$, $X \in \Pi$, then $\mu = \nu$ on $\sigma(\Pi)$.
Remark 1 - The intersection of an arbitrary family of $\lambda$-systems is a $\lambda$-system.
Remark 2 - If $\Lambda$ is both a $\Lambda$-system and a $\pi$-system, then it is a $\sigma$-algebra. (Intuition: "finite intersection + complements + increasing countable unions = countable unions")

Proof (of Theorem) - It's enough to show that $\lambda(\Pi) \supseteq \sigma(\Pi)$, since $\Lambda \supseteq \lambda(\pi)$. By Remark 2, it's enough to show that $\lambda(\pi)$ is a $\pi$-system, so then it's a $\sigma$-algebra.

For any $C \in \lambda(\pi)$, set $\Lambda_C = \{A \in \lambda(\pi) \mid A \cap C \in \lambda(\pi)\}$. To show that $\lambda(\pi)$ is a $\pi$-system, we show that $\Lambda_C = \lambda(\pi)$ for each $C \in \lambda(\pi)$. In particular, it's enough to show that for any $C \in \lambda(\pi)$, $\Lambda_C$ is a $\lambda$-system that contains $\pi$.

**Claim** $\Lambda_C$ is a $\lambda$-system.

**Proof** 1. Clearly $X \in \Lambda_C$.

2. If $A, B \in \Lambda_C$ with $A \subseteq B$, then

$$(B \backslash A) \cap C = (B \cap C) \backslash (A \cap C) \in \lambda(\pi)$$

since $A \cap C, B \cap C \in \lambda(\pi)$.

3. Let $(A_n) \subseteq \Lambda_C$ be increasing. Then

$$(\bigcup_{n \geqslant 1} A_n) \cap C = \bigcup_{n \geqslant 1} (A_n \cap C) \in \lambda(\pi)$$

since $A_n \cap C \subseteq A_{n+1} \cap C$. □

**Claim** $\Lambda_C \supseteq \Pi$.

**Proof** We need to show for all $B \in \Pi$, $B \cap C \in \lambda(\Pi)$. If $C \in \Pi$, this is immediate since $\Pi$ is a $\pi$-system. If $C \in \lambda(\Pi)$, fix $B \in \Pi$. By the previous case, $\Lambda_B \supseteq \lambda(\pi)$, so $B \cap C \in \lambda(\Pi)$, so $B \in \Lambda_C$. Thus, $\Lambda_C \supseteq \lambda(\pi)$, so we are done. □

□

**Lecture 6 - 09/08/2014**

**Regularity**

Definition - Let $(X, d)$ be a metric space, $\mathcal{B}(X)$ the Borel $\sigma$-algebra of $X$. We say $\mu$ is a *regular Borel measure* on $X$ if

1. $\mu$ is a measure on $(X, \mathcal{B}(X))$
2. For every $A \in \mathcal{B}(X)$, $\mu(A) = \inf\{\mu(U) \mid U \supseteq A \text{ is open}\}$
3. For every open $U \subseteq X$, $\mu(U) = \sup\{\mu(K) \mid K \subseteq U \text{ is compact}\}$
4. For every compact $K \in \mathcal{B}(X)$, $\mu(K) < \infty$.

A measure satisfying condition 2 is called an *outer regular measure*. A measure satisfying condition 3 is called an *inner regular measure*. A measure satisfying condition 4 is called a *Radon measure*.

If $A \in \mathcal{B}(X)$ satisfies condition 2, we say $A$ is *inner regular with respect to* $\mu$ or $\mu$*-inner regular*. If $A$ satisfies condition 3, we say $A$ is *outer regular with respect to* $\mu$ or $\mu$*-outer regular*.

Goal: Regularity of $\lambda$.

Theorem (Regular of Finite Borel Measures) - Let $X$ be a compact space. Let $\mu$ be a finite Borel measure on $X$. Then $\mu$ is regular.
Proof - Note, $\mu$ is Radon since it is finite. Define

$$\Lambda = \{A \in \mathcal{B}(X) \mid A \text{ is } \mu\text{-inner and } \mu\text{-outer regular}\}$$

It suffices to show that $\Lambda$ contains all open sets and that $\Lambda$ is a $\lambda$-system, since then $\Lambda \supseteq \mathcal{B}(X)$ (follows because open sets form a $\pi$-system).

Let $U \subseteq X$ be open. Then $U$ is trivially $\mu$-outer regular. For each $n$, define $K_n = \{x \in U \mid;\ d(x, U^c) \geqslant \frac{1}{n}\}$. Then clearly $K_n$ is closed. Since $X$ is compact, then $K_n$ is compact. Note, since $U$ is open, then $x \in U$ if and only if $d(x, U) > 0$. Therefore, $U = \bigcup_{n \geqslant 1} K_n$. Since $K_n \subseteq K_{n+1}$, then $\lim_{n\to\infty} \mu(K_n) = \mu(U)$, thus $U$ is $\mu$-inner regular, so $\Lambda$ contains all open sets.

We now argue that $\Lambda$ is a $\lambda$-system. Clearly $X \in \Lambda$ since $X$ is both open and compact. Now let $A_1, A_2 \in \Lambda$ with $A_1 \subseteq A_2$. Fix $\epsilon > 0$. Since $\mu$ is finite, then for each $i = 1, 2$, there are open $U_i$ and compact $K_i$ such that

$$K_i \subseteq A_i \subseteq U_i, \quad \mu(A_i \backslash K_i) < \epsilon, \quad \mu(U_i \backslash A_i) < \epsilon$$

Note that $K_2 \backslash U_1 \subseteq A_2 \backslash A_1 \subseteq U_2 \backslash K_1$, where $K_2 \backslash U_1$ is compact and $U_2 \backslash K_1$ is open. Moreover,

$$\mu((U_2 \backslash K_1) \backslash (A_2 \backslash A_1)) = \mu((U_2 \backslash A_2) \cup (A_1 \backslash K_1)) = \mu(U_2 \backslash A_2) + \mu(A_1 \backslash K_1) < 2\epsilon$$

and

$$\begin{aligned}
\mu((A_2 \backslash A_1) \backslash (K_2 \backslash U_1)) &= \mu((A_2 \cap A_1^c) \cap (K_2 \cap U_1^c)^c) \\
&= \mu((A_2 \cap A_1^c) \cap (K_2^c \cup U_1)) \\
&= \mu((A_2 \cap A_1^c \cap K_2^c) \cup (A_2 \cap A_1^c \cap U_1)) \\
&\leqslant \mu((A_1 \cap K_1^c) \cup (A_1^c \cap U_1)) \\
&\leqslant \mu(A_1 \backslash K_1) + \mu(U_1 \backslash A_1) \\
&< 2\epsilon
\end{aligned}$$

Letting $\epsilon \to 0^+$, we have $A_2 \backslash A_1 \in \Lambda$.

Finally, let $(A_n) \subseteq \Lambda$ be increasing. Fix $\epsilon > 0$. For each $n$, let $K_n \subseteq A_n \subseteq U_n$ be compact and open with $\mu(A_n \backslash K_n) < \frac{\epsilon}{2^n}$ and $\mu(U_n \backslash A_n) < \frac{\epsilon}{2^n}$. Denote $A = \bigcup\limits_{n \geqslant 1} A_n$ and $U = \bigcup\limits_{n \geqslant 1} U_n$. Then $A \subseteq U$, $U$ is open, and

$$\mu(U \backslash A) = \mu(\bigcup_{n \geqslant 1} (U_n \backslash A)) \leqslant \sum_{n \geqslant 1} \mu(U_n \backslash A) \leqslant \sum_{n \geqslant 1} \mu(U_n - A_n) < \epsilon$$

Let $E_n = \bigcup\limits_{1 \leqslant i \leqslant n} K_i$, $K = \bigcup\limits_{i \geqslant 1} K_i$. Note, each $E_n$ is compact (since it is closed) and $E_n \subseteq E_{n+1} \subseteq K$. Moreover, $\lim\limits_{n \to \infty} \mu(E_n) = \mu(K) < \infty$, so there is $N$ such that $\mu(E_N) > \mu(K) - \epsilon$, or $\mu(K \backslash E_N) < \epsilon$. Now $E_N \subseteq A$ and

$$\mu(A \backslash E_N) = \mu(A \backslash K) + \mu(K \backslash E_N) < \mu(\bigcup_{n \geqslant 1} (A_n \backslash K)) + \epsilon \leqslant \sum_{n \geqslant 1} \mu(A_n \backslash K) + \epsilon \leqslant \sum_{n \geqslant 1} \mu(A_n \backslash K_n) + \epsilon \leqslant 2\epsilon,$$

Letting $\epsilon \to 0^+$, we have $A \in \Lambda$, so $\Lambda$ is a $\lambda$-system. This finishes the proof. $\square$

**Lecture 7 - 09/10/2014**

Regularity of Lebesgue Measure

Proposition I - For all $A \in \mathcal{L}(\mathbb{R}^d)$, we have

$$\lambda(A) = \inf\{\lambda(U); U \supseteq A,\ U \text{ open}\} = \sup\{\lambda(K) \mid K \subseteq A,\ K \text{ compact}\}$$

and also $\lambda(K) < \infty$ for all compact $K$.
Proof - Clearly $\lambda(A) \leqslant \inf\limits_{U \supseteq A} \lambda(U)$. Moreover, countable subadditivity implies

$$\begin{aligned}\lambda(A) = \lambda^*(A) &= \inf\{\sum \ell(I_n) \mid I_n \text{ open cells},\ A \subseteq \bigcup_n I_n\} \\ &\geqslant \inf\{\lambda(\bigcup_n I_n) \mid I_n \text{ open cells},\ A \subseteq \bigcup_n I_n\} \\ &= \inf\{\lambda(U) \mid U \supseteq A,\ U \text{ open}\}\end{aligned}$$

Thus, $\lambda$ is outer regular. To show that $\lambda$ is inner regular, we reduce to the case that $A$ is bounded. (For the general case, consider the bounded sets $A \cap [-n, n]^d$.)

Since $A$ is bounded, there is a closed cell $I$ with $I \supseteq A$. Fix $\epsilon > 0$. By outer regularity, there is open $U \supseteq I \backslash A$ such that

$$\lambda(I \backslash A) \leqslant \lambda(U) \leqslant \lambda(I \backslash A) + \epsilon$$

Rearranging, we have

$$\lambda(A) \geqslant \lambda(I) - \lambda(U) = \lambda(I \backslash U) \geqslant \lambda(A) - \epsilon$$

We are done, since $I \backslash U$ is compact (closed and bounded).

Finally, note that compact sets clearly have finite measure because they are bounded. □

Proposition II - For all $A \in \mathcal{L}$, there is $\epsilon > 0$, $U$ open, and $C$ closed such that

$$C \subseteq A \subseteq U, \qquad \lambda(U \backslash C) < \epsilon$$

Proof - Case I: Suppose that $\lambda(A) < \infty$. By Proposition I, given $\epsilon > 0$ there are compact $K \subseteq A$ and open $U \supseteq A$ with

$$\lambda(A) - \tfrac{\epsilon}{2} < \lambda(K) \leqslant \lambda(A) \leqslant \lambda(U) < \lambda(A) + \tfrac{\epsilon}{2}$$

Now $\lambda(U \backslash K) < \epsilon$ and $K$ is closed, so we are done.

Case II: $A$ is arbitrary. Set $D_n = B(0, n+1) \backslash \overline{B(0, n)}$ and write $A = \bigcup\limits_n (A \cap D_n)$. By Case I, for each $n$ and $\epsilon > 0$, there are open $U_n$ and closed $C_n$ with

$$C_n \subseteq A \cap D_n \subseteq U_n, \qquad \lambda(U_n \backslash C_n) < \tfrac{\epsilon}{2^n}$$

Set $U = \bigcup\limits_n U_n$, $C = \bigcup\limits_n C_n$. Clearly $U$ is open. Moreover, $C$ is closed since $C_n \subseteq D_n$ for each $n$. Finally,

$$\lambda(U \backslash C) \leqslant \sum_n \tfrac{\epsilon}{2^n} = \epsilon$$

□

Corollary - Let $A \subseteq \mathbb{R}^d$. Then

$$A \in \mathcal{L} \iff A = F_\sigma \cup N,$$

where $F_\sigma$ is a countable union of closed sets and $\lambda(N) = 0$.

Existence of a Nonmeasurable Set

Proposition - $\mathcal{L}(\mathbb{R}) \subsetneq \mathcal{P}(\mathbb{R})$
Proof - Define an equivalence relation $\sim$ on $\mathbb{R}$ by $x \sim y$ if and only if $x - y \in \mathbb{Q}$. Partition $\mathbb{R}$ into equivalence classes, say for $\alpha \in \mathbb{R}$ we write

$$C_\alpha = \{x \in \mathbb{R} \mid x - \alpha \in \mathbb{Q}\} = \alpha/\sim$$

Let $A \subseteq \mathbb{R}$ be a set of representatives for the equivalence classes, i.e. for each $\alpha$, $A \cap C_\alpha$ is a singleton.

**Claim** $A \notin \mathcal{L}(\mathbb{R})$

**Proof** Suppose not, i.e. $A \in \mathcal{L}$. Then $\mathbb{R} = \bigcup_{q \in \mathbb{Q}} (A + q)$ (disjoint union), so

$$\lambda(\mathbb{R}) = \sum_{q \in \mathbb{Q}} \lambda(A + q) = \sum_{q \in \mathbb{Q}} \lambda(A) \implies \lambda(A) > 0$$

We argue that $\lambda(A) = 0$, which is a contradiction. By inner regularity, it suffices to show that $\lambda(K) = 0$ for all compact $K \subseteq A$. Given such $K$, set $E = \bigcup_{q \in \mathbb{Q}, |q| \leqslant 1} (K+q)$. Now that $E \in \mathcal{L}$ and $\lambda(E) < \infty$ since $E$ is bounded. Now

$$\lambda(E) = \sum_{q \in \mathbb{Q}, |q| \leqslant 1} \lambda(K) \implies \lambda(K) = \lambda(E) = 0$$

□

□

Remark - There is $A \subseteq \mathbb{R}$ such that

$$E \subseteq A,\ E \in \mathcal{L} \implies \lambda(E) = 0, \qquad E \subseteq A^c,\ E \in \mathcal{L} \implies \lambda(E) = 0$$

**Lecture 8 - 09/12/2014**

Completion

Proposition - Let $A \subseteq \mathbb{R}^d$. Then

$$A \in \mathcal{L}(\mathbb{R}^d) \iff \exists F, G \in \mathcal{B}(\mathbb{R}^d),\ F \subseteq A \subseteq G,\ \lambda(G \backslash F) = 0$$

Proof ( $\implies$ ) Suppose $A \in \mathcal{L}$. Then for each $n \in \mathbb{N}$, there are closed $C_n$, open $U_n$ with

$$C_n \subseteq A \subseteq U_n, \qquad \lambda(U_n \backslash C_n) < \tfrac{1}{n}$$

Let $F = \bigcup_n C_n$, $G = \bigcup_n U_n$. Now $\lambda(G \backslash F) \leqslant \lambda(U_n \backslash C_n) < \frac{1}{n}$ for all $n$, so $\lambda(G \backslash F) = 0$.

( $\impliedby$ ) Write $A = F \cup (A \backslash F)$. Since $A \backslash F \subseteq G \backslash F$ and $\lambda(G \backslash F) = 0$, then $A \backslash F \in \mathcal{L}$, so $A \in \mathcal{L}$. □

Let $(X, \Sigma, \mu)$ be a measure space.
Definition - Define $\mathcal{N} = \{A \subseteq X \mid \exists B \in \Sigma,\ B \supseteq A,\ \mu(B) = 0\}$, the set of all $\mu$-null subsets of $X$.
Definition - We say $\Sigma$ is *complete* with respect to $\mu$ if $\Sigma \supseteq \mathcal{N}$.
Remark - The $\sigma$-algebra of a measure constructed via an outer measure using the Carathéodory construction is complete.
Proof - Let $\mu$ be such a measure and let $N$ be any $\mu$-null set, say $M \supseteq N$ with $\mu(M) = 0$. For any $A \in X$, we have

$$\mu(A) \geqslant \mu(M) + \mu(A \cap N^c) \geqslant \mu(A \cap N) + \mu(A \cap N^c),$$

so $N \in \Sigma$. □
Corollary - $\mathcal{L}(\mathbb{R}^d)$ is complete.

Definition - We define the *completion* $\Sigma_\mu$ of $\Sigma$ with respect to $\mu$ by

$$\Sigma_\mu = \{A \cup N \mid A \in \Sigma,\ N \in \mathcal{N}\}$$

Definition - Given $A \in \Sigma_\mu$, define $\overline{\mu}(A) = \mu(B)$ if $A = B \cup N$, where $B \in \Sigma$ and $N \in \mathcal{N}$.
Remark - $\overline{\mu}$ is well-defined. (We prove this below.)
Lemma - $A \in \Sigma_\mu$ if and only if there are $F, G \in \Sigma$ such that $F \subseteq A \subseteq G$ and $\mu(G \backslash F) = 0$.
Proof - Let $A \in \Sigma_\mu$, say $A = B \cup N$ where $B \in \Sigma$, $N \in \mathcal{N}$. Since $N$ is $\mu$-null, there is $M \in \Sigma$ with $M \supseteq N$ and $\mu(M) = 0$. Let $F = B$, $G = B \cup M$. Then $F \subseteq A \subseteq G$, $\mu(G \backslash F) \leqslant \mu(M) = 0$.

Conversely, let $F, G \in \Sigma$ with $F \subseteq A \subseteq G$, $\mu(G \backslash F) = 0$. Write $A = F \cup (A \backslash F)$. Clearly $A \backslash F$ is $\mu$-null, so $A \in \Sigma_\mu$. □

Proposition - Let $(X, \Sigma, \mu)$ be a measure space, $(X, \Sigma_\mu, \overline{\mu})$ its completion.

1. $\Sigma_\mu$ is a $\sigma$-algebra.
2. $\overline{\mu}$ is a measure on $\Sigma_\mu$ and $\overline{\mu} \restriction_\Sigma = \mu$.
3. $\Sigma_\mu$ is complete with respect to $\overline{\mu}$.

Proof (1) Clearly $\varnothing \in \Sigma_\mu$. Given $A \in \Sigma_\mu$, the previous lemma implies there are $F, G \in \Sigma$ with $F \subseteq A \subseteq G$ and $\mu(G \backslash F) = 0$. Now $F^c, G^c \in \Sigma$ with $G^c \subseteq A^c \subseteq F^c$ and $\mu(F^c \backslash G^c) = \mu(G \backslash F) = 0$. Finally, let $(A_n) \subseteq \Sigma_\mu$, say $A_n = B_n \cup N_n$, where $B_n \in \Sigma$ and $N_n$ is $\mu$-null. Set $B = \bigcup_n B_n$, $N = \bigcup_n N_n$. Now $B \in \Sigma$ and by subadditivity, $N$ is $\mu$-null. Thus $\bigcup_n A_n \in \Sigma_\mu$.

(2) To see that $\overline{\mu}$ is well-defined, let $A \in \Sigma_\mu$, say $A = B_1 \cup N_1 = B_2 \cup N_2$, where $B_1, B_2 \in \Sigma$ and $N_1, N_2$ are $\mu$-null. Let $M_i \supseteq N_i$ for each $i$, $\mu(M_i) = 0$. Now

$$\mu(B_1) \leqslant \mu(B_2 \cup M_2) \leqslant \mu(B_2) + \mu(M_2) = \mu(B_2)$$

and similarly, $\mu(B_2) \leqslant \mu(B_1)$, so $\mu(B_1) = \mu(B_2)$. Thus, $\overline{\mu}$ is well-defined.

Clearly $\overline{\mu}(\varnothing) = 0$. Now let $(A_n) \subseteq \Sigma_\mu$ be disjoint, say $A_n = B_n \cup N_n$, where $B_n \in \Sigma$ and $N_n$ is $\mu$-null. Write $\bigcup\limits_n A_n = B \cup N$, where $B = \bigcup\limits_n B_n \in \Sigma$ and $N = \bigcup\limits_n N_n$ is $\mu$-null. Now

$$\overline{\mu}(\bigcup_n A_n) = \mu(B) = \sum_n \mu(B_n) = \sum_n \overline{\mu}(A_n),$$

since the sets $B_n$ are disjoint. Thus, $\overline{\mu}$ is a measure, and clearly $\overline{\mu} \restriction_\Sigma = \mu$.

(3) Let $N$ be null with respect to $\Sigma_\mu$, say $A \in \Sigma_\mu$ with $A \supseteq N$, $\overline{\mu}(A) = 0$. Let $F \subseteq A \subseteq G$ with $\mu(G \backslash F) = 0$, $F, G \in \Sigma$. Now $\mu(F) = \mu(A) = \mu(G) = 0$. In particular, we have $\varnothing \subseteq N \subseteq G$, $\mu(G \backslash \varnothing) = \mu(G) = 0$, so $N \in \Sigma_\mu$. Thus, $\Sigma_\mu$ is complete. $\square$

Remark - Taking the completion does *not* simply correspond to throwing in all sets of measure zero!

Example - $\mu = \lambda$, $X = \mathbb{R}^d$, $\Sigma = \{X, \varnothing\}$. Then $\Sigma_\mu = \{X, \varnothing\}$.

**Lecture 9 - 09/15/2014**

Goal: Measurable Functions

Definition - Let $(X, \Sigma)$ be a measure space, $(Y, \tau)$ a topological space (usually $\mathbb{R}$ or $[-\infty, \infty]$). We say $f : X \to Y$ is *measurable* if $f^{-1}(\tau) \subseteq \Sigma$, i.e. the preimage of an open set is measurable.

(Remark - The topology on $[-\infty, \infty]$ is generated by intervals $[-\infty, a)$, $(b, \infty]$.)

Example - If $X$ is a metric space, with $\Sigma \supseteq \mathcal{B}(X)$, then continuous functions are measurable.

Proposition - Let $\mathcal{B}(Y)$ be the Borel $\sigma$-algebra on $Y$. Then $f : X \to Y$ is measurable if and only if $f^{-1}(\mathcal{B}(Y)) \subseteq \Sigma$.

Proof ( $\Longleftarrow$ ) This is clear, since $f^{-1}(\tau) \subseteq f^{-1}(\mathcal{B}(Y))$.

( $\Longrightarrow$ ) We prove a useful lemma.

> **Lemma** Let $f : X \to Y$ be arbitrary. Then $\Sigma' = \{A \subseteq Y \mid f^{-1}(A) \in \Sigma\}$ is a $\sigma$-algebra.
>
> **Proof** Exercise. □

By the lemma, if $f : X \to Y$ is measurable, then $\Sigma'$ is a $\sigma$-algebra containing all open sets, thus it contains $\mathcal{B}(Y)$. □

Proposition - Let $f : X \to [-\infty, \infty]$. Then

$$\begin{aligned} f \text{ is measurable} &\iff \forall a \in \mathbb{R},\ f^{-1}([-\infty, a)) \in \Sigma \\ &\iff \forall a \in \mathbb{R},\ f^{-1}((a, \infty]) \in \Sigma \\ &\iff \forall a \in \mathbb{R},\ f^{-1}([a, \infty]) \in \Sigma \end{aligned}$$

Proof Clearly measurability implies the other three conditions by the previous lemma. Since the half-infinite intervals generate the Borel $\sigma$-algebra, the reverse implication is clear, since $\Sigma' = \{A \subseteq [-\infty, \infty] \,\|; f^{-1}(A) \in \Sigma\}$ is a $\sigma$-algebra. □

Proposition - Say $f : X \to \mathbb{R}^d$ is measurable, $g : \mathbb{R}^d \to \mathbb{R}^d$ is *Borel measurable* (i.e. $g^{-1}(\mathbb{R}^d) \subseteq \mathcal{B}(\mathbb{R}^d)$). Then $g \circ f$ is measurable.

Proof We have

$$(g \circ f)^{-1}(\mathcal{B}(\mathbb{R}^d) \subseteq f^{-1}(g^{-1}(\mathcal{B}(\mathbb{R}^d))) \subseteq f^{-1}(\mathcal{B}(\mathbb{R}^d) \subseteq \Sigma$$

□

Proposition - Let $f, g : X \to [-\infty, \infty]$ be arbitrary. Then

$$f, g \text{ are measurable} \iff (f, g) : X \to [-\infty, \infty]^2 \text{ is measurable}$$

Proof - Supppose that $f, g$ are measurable. Then for any open $U, V \subseteq [-\infty, \infty]$, we have

$$(f, g)^{-1}(U \times V) = f^{-1}(U) \cap g^{-1}(V) \in \Sigma$$

In particular, $(f, g)^{-1}(I) \in \Sigma$ for any open cell $I \subseteq [-\infty, \infty]$. Since the cells generate the Borel $\sigma$-algebra, then $(f, g)(\mathcal{B}([-\infty, \infty]^2) \subseteq \Sigma$, so $(f, g)$ is measurable.

Conversely, if $(f, g)$ is measurable, then so are $f = \pi_1 \circ (f, g)$ and $g = \pi_2 \circ (f, g)$ (here, $\pi_i$ denotes projection onto the $i$th coordinate, which is continuous). □

Corollary - Let $f, g : X \to \mathbb{R}$ be measurable. Then $f + g$ is measurable.

Proof - Denote $F(x, y) = x + y$, which is continuous, $G(x) = (f(x), g(x))$. Then $f + g = F \circ G$, which is continuous by the previous propositions. □

Proposition - Let $(f_n)$ be a sequence of measurable functions, $f_n : X \to [-\infty, \infty]$. Then the following functions are measurable.

1. $\sup f_n$
2. $\limsup f_n$
3. $\inf f_n$
4. $\liminf f_n$

Moreover: (5) if $f_n \to f$ pointwise, then $f$ is measurable.
Proof - Denote $g = \inf f_n$. Then $\{g \geqslant \alpha\} = \bigcup_n \{f_n \geqslant \alpha\}$, so since the $f_n$'s are measurable, so is $g$. Now $\sup f_n = -\inf(-f_n)$, so $\sup f_n$ is measurable. Combining (1) and (3), we get (2) and (4). Now either of (2) or (4) implies (5). $\square$

**Lecture 10 - 09/17/2014**

The Devil's Staircase

Recall - The Cantor set $C \subseteq [0,1]$ is compact, has Hausdorff dimension $\alpha = \frac{\ln 2}{\ln 3}$, and $H^\alpha$ denotes the Hausdorff measure of dimension $\alpha$.

Definition - The *Cantor function* (also known as devil's staircase) is defined by $f(x) = \frac{H^\alpha(C \cap [0,x])}{H^\alpha(C)}$. Intuitively, this is the fraction of the Cantor set that lies to the left of $x$.
Remarks

1. $f$ is continuous, since $|f(y) - f(x)| \leqslant C_i|x-y|^\alpha$ for some $C_i \geqslant 0$ and $\alpha \in (0,1)$ (exercise).
2. $f$ is increasing (since $H^\alpha$ is a measure).

Set $g(x) = \inf\{f = x\} = \inf\{y \mid f(y) = x\}$. Note, since $f$ is continuous, then the infimum is attained, so $f(g(x)) = x$. However, it's not hard to see that $g$ isn't continuous since $g(\frac{1}{2}) = \frac{1}{3}$, while $g(y) \geqslant \frac{2}{3}$ for any $y > \frac{1}{2}$. Note that

1. For all $x \in [0,1]$, $g(x) \in C$ (exercise).
2. $g$ is strictly increasing, since $x < y \iff f(g(x)) < f(g(y)) \implies g(x) < g(y)$.
3. $g$ is Borel measurable, since

$$\begin{aligned}\{g(x) \leqslant \alpha\} &= \{x \mid g(x) \leqslant \alpha\} \\ &= \{x \mid x = f(g(x)) \leqslant f(\alpha)\} \quad (f \text{ is increasing}) \\ &= \{x \mid x \leqslant f(\alpha)\}\end{aligned}$$

Proposition - $\mathcal{L}(\mathbb{R}) \supsetneq \mathcal{B}(\mathbb{R})$.
Proof - Let $A \subseteq [0,1]$, $A \notin \mathcal{L}$ (we know such $A$ exists). Set $B = g(A)$. Since $B \subseteq C$ and $\lambda(C) = 0$, then $B \in \mathcal{L}$ by completeness. On the other hand, $B \notin \mathcal{B}$, since $g^{-1}(B) = A \notin \mathcal{B}$ and $g$ is Borel measurable. $\square$

Remark - Let $h_1, h_2 : \mathbb{R} \to \mathbb{R}$, $h_1$ Lebesgue measurable and $h_2$ Borel measurable. Then $h_1 \circ h_2$ need not be Lebesgue measurable.
Proof - Let $h_1 = \chi_B$ ($B$ as above), $h_2 = g$. Then

$$h_1 \circ h_2 = \chi_B \circ g = \chi_{g^{-1}(B)} = \chi_A,$$

which is not measurable since $A$ is not measurable. $\square$

Definition - Let $(\Sigma, X, \mu)$ be a measure space. We say property $P$ holds *almost everywhere* (abbreviated a.e.) on some set $A$ if there is a null set $N$ such that $P$ holds on $A \backslash N$. We say $P$ holds a.e. if $P$ holds a.e. on $X$.
Notation - Some say $\tilde{\forall} x$ instead of a.e. x. We won't use this notation.

Examples - Almost every real is irrational. The function $f(x) = |x|$ is differentiable a.e.

Theorem (Rademacher) - Any Lipschitz function $f : \mathbb{R} \to \mathbb{R}$ is differentiable almost everywhere.

Example - Continued Fractions - Given $x \in [0,1]$, one may write $x = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\cdots}}}$, where $a_i \in \mathbb{Z}$ is uniquely determined by $x$, say $a_i = a_i(x)$. Define $\frac{p_n(x)}{q_n(x)} = \frac{1}{a_1 + \frac{1}{\cdots \frac{1}{a_n + 1}}}$ in reduced terms. We know that $\frac{p_n(x)}{q_n(x)} \to x$ as $n \to \infty$. Typically (when $x$ is irrational), $q_n(x) \to \infty$.

Theorem - $\lim\limits_{n\to\infty} \frac{\ln q_n(x)}{n}$ exists and is equal to $\frac{\pi^2}{12\ln(2)}$ almost everywhere.

Proof - Comes from ergodic theory. □

Proposition - Say $(X, \Sigma, \overline{\mu})$ is the completion of $(X, \Sigma, \mu)$. Then

$$f : X \to \mathbb{R} \text{ is } \Sigma_\mu\text{-measurable} \iff \exists g : X \to \mathbb{R},\ g \text{ is } \Sigma\text{-measurable},\ f = g \text{ a.e.}$$

Proof - ( $\Longleftarrow$ ) Given $g$, let $U \subseteq \mathbb{R}$ be open. Then

$$f^{-1}(U) = (\{f = g\} \cap g^{-1}(U)) \cup (\{f \neq g\} \cap f^{-1}(U))$$

Since the first set is in $\Sigma$ and the second is null, then $f^{-1}(U) \in \Sigma_\mu$.

( $\Longrightarrow$ ) If $f = \chi_E$, where $E \in \Sigma_\mu$, this is clear, since we may write $E = E' \cup N$, where $E' \in \Sigma$ and $N$ is null. Now $g = \chi_{E'}$ does the trick. Since the result holds for indicator functions, it holds for simple functions.

The result for measurable functions will follow from a theorem next time about approximating measurable functions by simple functions. □

**Lecture 11 - 09/19/2014**

Approximation of Measurable Functions

Definition - We say $s : X \to \mathbb{R}$ is *simple* if it is Borel measurable and has finite range.

Definition - Given $A \in \Sigma$, the *characteristic function of* $A$ is $\chi_A$, $\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$.

Remark - If $s$ has a finite range, then $s$ has the form

$$s = \sum_{1 \leqslant i \leqslant N} a_i s^{-1}(\{s_i\}),$$

where range$(s) = \{s_1, \ldots, s_N\}$. Hence, $s$ is simple if and only if $s^{-1}(\{s_i\})$ is measurable for each $i$.

Proposition 1 - If $f : X \to [0, \infty)$ is measurable, then there exists an increasing sequence $(s_n)$ of simple functions such that $s_n \nearrow f$ pointwise.

Proof - Define $A_{k,n} := f^{-1}([\frac{k}{2^n}, \frac{k+1}{2^n}))$ and let $s_n = \sum\limits_{1 \leqslant k \leqslant n2^n} \frac{k}{2^n} \chi_{A_{k,n}} + \chi_{\{f \geqslant n\}}$. Check that this works. □

Proposition 2 - If $f : X \to \mathbb{R}$ is measurable, then there exists a sequence $(s_n)$ of simple functions such that $|s_n| \leqslant f$ for all $n$ and $s_n \to f$ pointwise.

Proof - Denote $f^+ = \max(f, 0)$, $f^- = \max(-f, 0)$, so $f = f^+ - f^-$. Find $(s_n)$ and $(t_n)$ with $s_n \nearrow f^+$ and $t_n \nearrow f^-$, so now $s_n - t_n \to f$. Moreover,

$$|s_n - t_n| \leqslant |s_n| + |t_n| \leqslant f^+ + f^- = |f|$$

□

Theorem (Lusin's Theorem) - Let $X$ be a metric space, $\mu$ a finite regular measure. If $f : X \to \mathbb{R}$ is $\mu$-measurable, then for every $\epsilon > 0$, there is $g : X \to \mathbb{R}$ continuous such that $\mu(\{f \neq g\}) < \epsilon$.

Lemma 1 - For each $\epsilon > 0$, there is closed $C \subseteq X$ such that $\mu(X \backslash C) < \epsilon$ and $f \restriction_C : C \to \mathbb{R}$ is continuous.

Lemma 2 (Tietze Extension Theorem: Easy Version) - Let $C \subseteq X$ be closed, $h : C \to \mathbb{R}$ continuous. Then there is continuous $H : X \to \mathbb{R}$ such that $H \restriction_C = h$.

Proof (Lusin's Theorem) - Follows from Lemmas 1 and 2. □

Proof (Lemma 2) - Define $H(x) = \begin{cases} h(x) & x \in C \\ \inf\limits_{c \in C} \{h(c) + \frac{d(x,c)}{d(x,C)} - 1\} & x \notin C \end{cases}$. Not hard to check that this is continuous. Note, this only works if $h$ is bounded below on $C$. If $h$ is unbounded, then consider arctan$(h)$ instead. □

Proof (Lemma 1) - Case I: Suppose that $f$ is bounded. Without loss of generality, $f : X \to [0, 1]$. Set $A_{k,n} = f^{-1}([\frac{k}{2^n}, \frac{k+1}{2^n}))$, where $0 \leqslant k \leqslant 2^n$. For all $k, n$, there exist compact sets $K_{k,n} \subseteq A_{k,n}$ such that

$$\mu(A_{k,n} \backslash K_{k,n}) < \tfrac{\epsilon}{2^n(2^n+1)}$$

Now set $C_n = \bigcup\limits_{0 \leqslant k \leqslant 2^n} K_{k,n}$. Note that by construction, we have $\mu(C_n^c) < \frac{\epsilon}{2^n}$ by construction, since $\mu(C_n^c) \leqslant \sum\limits_k \mu(A_{k,n} \backslash K_{k,n})$. Now set $C = \bigcap\limits_n C_n$.

**Claim** This set $C$ works.

**Proof** Clearly $C$ is closed. Moreover, $\mu(C^c) \leqslant \sum_n \mu(C_n^c) = \epsilon$. To see that $f \restriction_C$ is continuous, denote $s_n = \sum_k \frac{k}{2^n}\chi_{A_{k,n}}$. Note that $s_n : C_n \to \mathbb{R}$ is continuous (by the pasting lemma for continuous functions) since $C_n = \bigcup_K K_{k,n}$ and $s_n \restriction_{K_{k,n}}$ is constant. Since $C \subseteq C_n$, then $s_n : C \to \mathbb{R}$ is continuous.

On $C$, $|f - s_n| \leqslant \frac{1}{2^n}$, so $s_n \to f$ uniformly. Thus, $f$ is continuous on $C$. □

□

Remark - This theorem is still true if $X = \mathbb{R}^d$, $\mu = \lambda$ because we only need the sets $K_{k,n}$ to be closed (not necessarily compact) such that $\lambda(A_{k,n} \backslash K_{k,n}) < \frac{\epsilon}{2^n(2^n+1)}$ and the rest of the proof is identical.

**Lecture 12 - 09/22/2014**

Integration

Definition - Let $s : X \to \mathbb{R}$ is simple, measurable, and nonnegative, say $s = \sum_i a_i \chi_{A_i}$, where $A_i = s^{-1}(a_i)$. Define

$$\int_X s \, d\mu = \sum_i a_i \mu(A_i)$$

with the convention $0 \cdot \infty = 0$.
Definition - Given $f : X \to [0, \infty]$ measurable, define

$$\int_X f \, d\mu := \sup_{\substack{0 \leqslant s \leqslant f \\ s \text{ simple meas.}}} \int_X s \, d\mu$$

Definition - Given $f : X \to [-\infty, \infty]$ measurable, define $f^+ = \max(f, 0)$, $f^- = \max(-f, 0)$ (note: these are measurable functions since $x \mapsto |x|$ is continuous and continuous$\circ$measurable is measurable).

1. If $\int_X f^+ \, d\mu$, $\int_X f^- \, d\mu < \infty$, we say $f$ is *integrable* (or $\mathcal{L}^1$) and define

$$\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu$$

2. If either $\int_X f^+ \, d\mu$ or $\int_X f^- \, d\mu$ is finite, we say $f$ is *integrable in the extended sense* and define

$$\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu \in [-\infty, \infty]$$

Notation - Depending on whether the measure or measure space are clear from context, we will sometimes write

$$\int_X f \, d\mu = \int f \, d\mu, \qquad \int_X f = \int f$$

Example 1 - Let $X = \mathbb{N}$, $\mu$ the counting measure. All sets are measurable, so all functions $f : X \to \mathbb{R}$ are measurable. Note a function $a : \mathbb{N} \to \mathbb{R}$ is just a sequence of real numbers. What is $\int_{\mathbb{N}} a \, d\mu$? Well, we'd like to say it's just $\sum_n a_n$. However, since the integral is defined as the positive part minus the negative part, we can only conclude that $\int_{\mathbb{N}} a \, d\mu = \sum_n a_n$ in the case that $\sum |a_n| < \infty$.
Remark - The above example illustrates that we've somehow lost cancellation. Even a conditionally convergent sum (such as $\sum_n \frac{(-1)^n}{n}$ won't have $\int_{\mathbb{N}} a \, d\mu$ defined, since we split up $a = a^+ - a^-$.

Proposition 1 - Let $s \geqslant 0$ be simple and measurable with $s(X) = \{a_1, \ldots, a_n\}$, $A_i = s^{-1}(a_i)$. Then

$$\sup_{\substack{0 \leqslant t \leqslant s \\ t \text{ simple meas.}}} \int_X t \, d\mu = \sum_i a_i \mu(A_i)$$

Proof - The left-hand side is clearly at least as large as the right-hand side. For the reverse inequality, it's enough to prove the following lemma.

**Lemma** Let $s$, $t$ be simple and measurable with $0 \leqslant s \leqslant t$. Then

$$\int_X s \, d\mu \leqslant \int_X t \, d\mu$$

**Proof** Exercise. □

□

Proposition 2 - If $f, g$ are integrable, then $\int_X (f+g)\, d\mu = \int_X f\, d\mu + \int_X g\, d\mu$.
Proof - We can't do this yet! Let's see what goes wrong. Start by assuming that $f, g \geqslant 0$. Let $0 \leqslant s \leqslant f$, $0 \leqslant t \leqslant g$, so that $s + t \leqslant f + g$. Now it's not hard to show that

$$\int_X s\, d\mu + \int_X t\, d\mu = \int_X (s+t)\, d\mu \leqslant \int_X (f+g)\, d\mu$$

Taking the supremum over $0 \leqslant s \leqslant f$, $0 \leqslant t \leqslant g$, we have

$$\int_X f\, d\mu + \int_X g\, d\mu \leqslant \int_X (f+g)\, d\mu$$

The problem is that we can't easily get the reverse! What to do? (To be continued)
Remark - To finish this off, we need some sort of limit theorem to take care of the sup in the definition of $\int_X f\, d\mu$.

Theorem (Lebesgue's Monotone Convergence Theorem) - Say $f_n : X \to [0, \infty]$ are measurable with $f_n \leqslant f_{n+1}$. Denote $f = \lim f_n$ (this exists because $f$ is monotone). Then

$$\lim \int_X f_n\, d\mu = \int_X f\, d\mu$$

Proof - For all $n$, we have $\int f_n \leqslant \int f_{n+1} \leqslant \int f$ (this follows from monotonicity of the integral. In particular, $\lim \int f_n$ exists and $\lim \int f_n \leqslant \int f$.

We need to show that $\lim \int f_n \geqslant \int f$. To do this, let $0 \leqslant s \leqslant f$ simple, measurable. We want $\lim \int f_n \geqslant \int s$. First, we take a naïve approach (which we'll fix in a moment). Write $s(X) = \{a_1, \dots, a_N\}$, $A_i = s^{-1}(a_i)$. Set $E_n = \{f_n \geqslant s\}$. Then $(E_n)$ is increasing and now

$$\int_X f_n\, d\mu \geqslant \int_X \chi_{E_n} f\, d\mu \geqslant \int_X \chi_{E_n} s\, d\mu = \sum_{i=1}^N a_i \mu(A_i \cap E_n)$$

We'd *like* to take the limit $n \to \infty$ and sy $\mu(A_i \cap E_n) \to \mu(A_i)$ to get the desired result. The problem is that this only works if $\bigcup_n E_n = X$. This may not happen if the sequence $(f_n)$ doesn't cross $s$ everywhere.

The solution is to instead let $0 < \epsilon < 1$ and set $E_n = \{f_n \geqslant (1-\epsilon)s\}$. Now $\bigcup_n E_n = X$. We modify the above inequality and get

$$\int_X f_n\, d\mu \geqslant \int_X \chi_{E_n} f\, d\mu \geqslant \int_X \chi_{E_n} (1-\epsilon) s\, d\mu = (1-\epsilon) \sum_{i=1}^N a_i \mu(A_i \cap E_n)$$

Taking the limit $n \to \infty$, we have $\lim \int f_n \geqslant (1-\epsilon) \int s$. Take $\epsilon \to 0^+$ to get $\lim \int f_n \geqslant \int s$, as desired. □

Proposition (Linearity) - If $f, g$ are integrable, then $f + g$ is integrable and $\int (f+g) = \int f + \int g$.
Proof - We do this in four steps.

(1) Let $s, t \geqslant 0$ be simple and integrable. Then $\int (s+t) = \int s + \int t$ (exercise).

(2) Now let $f, g \geqslant 0$ be measurable and integrable. Want $\int(f+g) = \int f + \int g$. Last time, we showed that there exist increasing sequences $(s_n)$, $(t_n)$ of nonnegative measurable simple functions with $s_n \to f$ and $t_n \to g$ pointwise. By MCT,

$$\lim \int s_n = f, \quad \lim \int t_n = g \implies \int (f+g) = \lim \int (s_n + t_n) = \lim(\int s_n + \int t_n) = \int f + \int g$$

(3) Lemma - Let $f$ be integrable, $f = g - h$, where $g, h \geqslant 0$ are integrable. Then

$$\int f = \int g - \int h$$

Proof - Write $f = f^+ - f^- = g - h$. Now

$$f^+ + h = g + f^- \implies \int f^+ + \int h = \int g + \int f^- \implies \int f = \int (f^+ - f^-) = \int g - \int h$$

□

(4) Now let $f, g$ be arbitrary and integrable, say $f = f^+ - f^-$, $g = g^+ - g^-$. Now

$$(f+g) = (f^+ + g^+) - (f^- + g^-)$$

Applying the lemma from (3), we have

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

□

**Lecture 13 - 09/24/2014**

Last time, we defined the integral, proved the Monotone Convergence Theorem, and used it to prove linearity.

Quote - "The selling point of the Lebesgue integral is the interchangeability of $\int$ and lim."

Remark - Note that the MCT fails if we don't require $f_n \geqslant 0$. Even if we assume $f_n \geqslant -1$, for example, consider the case $X = \mathbb{R}$, $f_n = -\frac{1}{n}$. Now $\int f_n = -\infty$ for all $n$, but $\int \lim f_n = 0$.

Goal - If $f_n \geqslant 0$ and $\lim f_n$ exists, want to say $\lim \int f_n = \int \lim f_n$.

Two Counterexamples

1. Mass escaping to infinity: Set $f_n = \chi_{[n,n+1)}$. Then
$$\lim \int f_n = 1 \neq 0 = \int \lim f_n$$
2. Mass collects at a point: Set $f_n = n\chi_{(0,\frac{1}{n}]}$. Then
$$\lim \int f_n = 1 \neq 0 = \int \lim f_n$$

To see what goes wrong here, consider the following theorem.

Theorem (Lebesgue Dominated Convergence Theorem) - Suppose that

1. $f_n : X \to [0, \infty]$ is measurable.
2. $f_n \to f$ pointwise.
3. There is integrable $F$ such that $|f_n| \leqslant F$ for all $n$.

Then $\lim \int f_n = \int f$.
Remark - This addresses the counterexamples above. In the first example, the 'envelope' of $(f_n)$ is the function $f(x) = 1$, which is not integrable. In the second example, the envelope of $(f_n)$ is $f(x) = \frac{1}{x}$, also not integrable.

To prove LDCT, we first need a lemma.

Fatou's Lemma - Let $f_n \geqslant 0$ be measurable, $f_n \to f$ pointwise. Then
$$\liminf \int f_n \geqslant \int f$$

Quote - "Mass can escape, but cannot be created."Proof - Set $g_k = \inf\{f_n \mid k \geqslant n\}$. Note that

$g_k \to f$ and $(g_k)$ is increasing, nonnnegative, so by MCT,
$$\lim \int g_k \geqslant \int f$$

Now $f_n \geqslant g_n$ for all $n$, so $\int f_n \geqslant \int g_n$ by monotonicity. Thus,

$$\liminf \int f_n \geqslant \liminf \int g_n = \lim \int g_n = \int f$$

$\square$

Remark - We cannot conclude that $\lim \int f_n \geqslant \int f$, since $\lim \int f_n$ may not exist.

Proof (LDCT) - Set $g_n = f + f_n$. Now $g_n \geqslant 0$, so Fatou's lemma implies

$$\int F + \liminf \int f_n = \liminf \int g_n \geqslant \int (F + f) = \int F + \int f$$

Cancelling $\int F$, we have $\liminf \int f_n \geqslant \int f$.

Now set $h_n = F - f_n$. Since $h_n \geqslant 0$, Fatou's lemma implies

$$\int F - \limsup \int f_n = \liminf \int h_n \geqslant \int (F - f) = \int F - \int f$$

Rearranging, we have $\limsup \int f_n \leqslant \int f$. Now

$$\limsup \int f_n \leqslant \int f \leqslant \liminf \int f_n \implies \lim \int f_n = \int f$$

$\square$

Quote - "The only work we've done to prove this was in the Monotone Convergence Theorem, and even that was only a few lines. The key is that we'ved coded all these results into the definition of measure."

Application - The Laplace Transform

Notation - We denote $\int_X f(x)\, d\mu(x) = \int_X f\, d\mu$. For example, $\int_0^1 x^2\, d\lambda(x)$ is the integral of $x^2$ with respect to $\lambda$ on $[0, 1]$. We denote integration with respect to $\lambda$ by $d\lambda(x) = dx$. Why do we do this?
IOU - If $f : [0, 1] \to \mathbb{R}$ is Riemann integrable, then $f$ is $\mathcal{L}$-integrable. Moreover, the Riemann integral and Lebesgue integral agree with one another.
Remark - An *improper* Riemann integral may disagree with the Lebesgue integral, since the Lebesgue integral doesn't allow for cancellation of an infinite positive part and an infinite negative part.

Suppose now that $f : [0, \infty) \to \mathbb{R}$ is integrable. Define the Laplace transform of $f$ by

$$F(s) = \int_0^\infty e^{-st} f(t)\, dt$$

Note: $F : (0, \infty) \to \mathbb{R}$. Here's an example of where Lebesgue integration is helpful.

Proposition

1. If $\int_0^\infty |f| < \infty$, then $F$ is continuous.
2. If $\int_0^\infty t|f(t)| < \infty$, then $F$ is differentiable (in fact, $C^1$).

Proof - For (1), given $s_0 \in (0, \infty)$, we have $\lim\limits_{s \to s_0} = \lim\limits_{s \to s_0} \int_0^\infty e^{-st} f(t)\, dt$. We want to apply LDCT. To get $F(s_0)$ on the right-hand side, let $(s_n)$ be a sequence with $s_n \to s_0$. Note that $|e^{-s_n t} f(t)| \leqslant |f(t)|$. Since $\int_0^\infty |f(t)|\, dt < \infty$, then by LDCT, we have

$$\lim_{n \to \infty} F(s_n) = F(s_0)$$

To prove (2), write $\frac{F(s)-F(s_0)}{s-s_0} = \int_0^\infty \frac{e^{-st}-e^{-s_0 t}}{s-s_0} f(t)\, dt$. For any $s \neq s_0$, the Mean Value Theorem implies that there is $s_1$ such that

$$\frac{e^{-st}-e^{s_0 t}}{s-s_0} = -te^{-s_1 t}$$

Thus, $|\frac{e^{-st}-e^{s_0 t}}{s-s_0}| \leqslant t|f(t)|$, so since $\int_0^\infty t|f(t)|\, dt < \infty$, then the LDCT implies that

$$\lim_{s \to s_0} \frac{F(s)-F(s_0)}{s-s_0} = \int_0^\infty -te^{-st} f(t)\, dt = \mathcal{L}(tf(t))$$

By (1), $\mathcal{L}(tf(t))$ is continuous, so we are done. □

**Lecture 14 - 09/26/2014**

We begin with an application of MCT.

Example (Beppo Levi's Theorem) - Let $(f_n)$ be a sequence of nonnegative measurable functions. Then

$$\int_X \sum_n f_n \, d\mu = \sum_n \int_X f_n \, d\mu$$

Proof - Set $s_N = \sum_{1 \leqslant n \leqslant N} f_n$. Note, $0 \leqslant s_n \leqslant s_{n+1}$. Now apply the MCT. □

Example - Let $\phi : \mathbb{R} \to [0, \infty]$ with $\int_{\mathbb{R}} \phi < \infty$. Define

$$f(x) = \sum_{m \in \mathbb{N}, n \in \mathbb{Z}^+} \phi((x - \tfrac{m}{n})2^{m+n})$$

Note that

1. $f = \infty$ at all rationals if $\phi(0) > 0$.
2. $f < \infty$ a.e. To see this, note that Beppo Levi implies

$$\int_{\mathbb{R}} f = \sum_{m,n} \int_{\mathbb{R}} \phi((x - \tfrac{m}{n})2^{m+n}) = \sum_{m,n} 2^{-(m+n)} \int_{\mathbb{R}} \phi < \infty$$

   In particular, $f < \infty$ a.e.

We'll spend most of the next week thoroughly studying convergence. Before that, let's mention a few things.

Pushforward Measures

Definition - Let $(X, \Sigma, \mu)$ be a measure space, $f : X \to Y$ arbitrary. Define $\Sigma' \subseteq \mathcal{P}(Y)$ by

$$\Sigma' = \{A \mid f^{-1}(A) \in \Sigma\}$$

Define $\nu$ on $\Sigma'$ by $\nu(A) = \mu(f^{-1}(A))$.
Proposition - $\nu$ is a measure on $(Y, \Sigma')$. Moreover, if $g : Y \to [-\infty, \infty]$ is integrable, then

$$\int_Y g \, d\nu = \int_X g \circ g \, d\mu$$

Proof - We've already seen that $\Sigma'$ is a $\sigma$-algebra. Moreover, it's easy to check that $\nu$ is a measure. To prove the statement about integrability, it's enough to show it for simple functions by the MCT and linearity. In fact, we may consider $s = \chi_A$, where $A \in \Sigma'$. Now

$$\int_Y s \, d\nu = \nu(A) = \mu(f^{-1}(A)) = \int_X \chi_{f^{-1}(A)} \, d\mu = \int_X s \circ f \, d\mu$$

Now $\int_Y g \, d\nu = \int_X g \circ f \, d\mu$ for all $g$ positive and simple. Thus, this is true for any integrable $g$. □

Application - Fix $x_0 \in \mathbb{R}^d$, $f : \mathbb{R}^d \to \mathbb{R}^d$ integrable. Then

$$\int_{\mathbb{R}^d} f(x) \, dx = \int_{\mathbb{R}^d} f(x + x_0) \, dx$$

Proof - Let $g(x) = x + x_0$. Equip the second $\mathbb{R}^d$ with the pushforward measure $\mu$ induced by $(\mathbb{R}^d, \lambda)$. Now $\mu(A) = \lambda(g^{-1}(A)) = \lambda(A)$, so by translation invariance, $\mu = \lambda$. The previous proposition now yields

$$\int_{\mathbb{R}^d} f(x)\, d\lambda(x) = \int_{\mathbb{R}^d} f(x)\, d\mu(x) = \int_{\mathbb{R}^d} f(x + x_0)\, d\lambda(x)$$

□

It's time for one last remark about integration.

Proposition - Let $(X, \Sigma, \mu)$ be complete and let $f : X \to \mathbb{R}$ be integrable, $f = g$ a.e. Then $g$ is integrable and $\int f = \int g$.
Proof - We already know $g$ is measurable. To prove that $g$ is integrable, set $h = f - g$. Then $h = 0$ a.e. and $h$ is measurable.

> **Claim** $\int h = 0$.
>
> **Proof** Decompose $h = h^+ - h^-$. We'll show that $\int_X h^+ = \int_X h^- = 0$. To do this, let $0 \leqslant s \leqslant h^+$ be simple, say $s = \sum a_i \chi_{A_i}$, where the sets $A_i$ are disjoint and measurable. If $a_i \neq 0$, then $A_i \subseteq \{h^+ > 0\}$, so $\mu(A_i) = 0$. Thus, $\int_X s\, d\mu = 0$, so now $\int_X h^+\, d\mu = 0$. Similarly, $\int_X h^- = 0$. This finishes the proof. □

Now $h$ integrable implies $f - h = g$ integrable, and $\int g = \int (f - h) = \int f$. □
Remark - In light of this remark, the MCT and LDCT (for $\mathbb{R}^d$) are still true if we replace $f_n \to f$ with $f_n \to f$ a.e.

Convergence

Theorem (Egorov's Theorem) - Let $(X, \Sigma, \mu)$ be a measure space, $\mu(X) < \infty$. Let $(f_n)$ be a sequence of measurable functions with $f_n \to f$ pointwise a.e. Then for any $\epsilon > 0$, there is $A \subseteq X$ such that $\mu(X \backslash A) < \epsilon$ and $f_n \to f$ uniformly on $A$.
Remark - In general, we can't do better, i.e. we can't get $\mu(A^c) = 0$.
Proof (of theorem) - For each $k$, there is $n_0$ such that $n > n_0$ implies $|f_n - f| < \frac{1}{k}$ a.e. This is because

$$\mu(\bigcup_m \bigcap_{n \geqslant m} \{|f_n - f| < \tfrac{1}{k}\}) = \mu(X)$$

In particular, for each $k$ there is $n_k$ such that

$$\mu(\bigcap_{n \geqslant n_k} \{|f_n - f| < \tfrac{1}{k}\}) \geqslant \mu(X) - \tfrac{\epsilon}{2^k}$$

Now set $A = \bigcap_k A_k$. Then

- $\mu(A) \geqslant \mu(X) - \epsilon$.
- Fix $k \in \mathbb{N}$. For all $a \in A$, we have $a \in A_k$, so $|f_n(a) - f(a)| < \frac{1}{k}$ for all $n \geqslant n_k$. Since $k$ was arbitrary, then $f_n \to f$ uniformly on $A$.

□

**Lecture 15 - 09/29/2014**

Definition (a.e. pointwise convergence) We say $f_n \to f$ $\mu$-a.e. if there exists a $\mu$-null set $N \in \Sigma$ such that $f_n \to f$ pointwise on $N^c$, i.e. if $\{f_n \nrightarrow f\}$ is null.
Definition (convergence in measure) We say $f_n \to f$ in $\mu$-measure (and write $f_n \xrightarrow{\mu} f$) if for every $\epsilon > 0$, we have

$$\lim_{n\to\infty} \mu(\{x \in X \mid |f_n - f| > \epsilon\}) = 0$$

Definition ($L^p$ convergence) We say $f_n \to f$ in $L^1$ if $\int_X |f_n - f| \, d\mu \to 0$ as $n \to \infty$. More generally, for $p \geqslant 1$, we say $f_n \to f$ in $L^p$ (and write $f_n \xrightarrow{L^p} f$) if

$$\lim_{n\to\infty} \int |f_n - f|^p = 0$$

Later we will define a similar notion for $p = \infty$.

Intuition - Think of $L^p$ as convergence in a metric space, where the metric is

$$d(f, g) = (\int_X |f - g|^p \, d\mu)^{1/p}$$

(We need $p \geqslant 1$, otherwise the triangle inequality fails.)

Example - To see that $f_n \to f$ a.e. does not imply $f_n \xrightarrow{\mu} f$, consider $f_n = \chi_{[n,n+1]}$. Then $f_n \to 0$ a.e. but $\mu(\{|f_n - 0| > \epsilon\}) = 1$ for all $n$ and $0 < \epsilon < 1$.

Proposition 1 - If $\mu(X) < \infty$ and $f_n \to f$ a.e., then $f_n \to f$ in measure.
Proof - Let $\epsilon > 0$. Since $f_n \to f$ a.e., then $\chi_{\{|f_n - f| > \epsilon\}} \to 0$ a.e. Moreover, $\chi_{\{|f_n - f| > \epsilon\}} \leqslant 1$, which is integrable since $\mu(X) < \infty$. By LDCT, we have

$$\lim_{n\to\infty} \mu(\{|f_n - f| > \epsilon\}) = \lim_{n\to\infty} \int \chi_{\{|f_n - f| > \epsilon\}} = 0$$

□

Remark - This can also be proved using Egorov's theorem.

Question - Does the converse hold? I.e. does $f_n \xrightarrow{\mu} f$ imply $f_n \to f$ a.e.?
Answer - No! For example, $f_1 = \chi_{[0,\frac{1}{2})}$, $f_2 = \chi_{[\frac{1}{2},1)}$, $f_3 = \chi_{[0,\frac{1}{4})}$, $f_4 = \chi_{[\frac{1}{4},\frac{1}{2})}$, etc. Set $f = 0$. Then clearly $\mu(\{|f_n - f| > \epsilon\}) \to 0$ for all $\epsilon > 0$, so $f_n \to f$ in measure. However, for all $x \in [0, 1)$, $f_n(x) = 0$ infinitely often and $f_n(x) = 1$ infinitely often, so $f_n$ fails to converge a.e. □

Proposition 2 - Suppose $f_n \to f$ in measure. Then there is a subsequence $(f_{n_k}) \subseteq (f_n)$ with $f_{n_k} \to f$ a.e.
Proof - Let $n_1$ be such that $\mu(\{|f_n - f| > 1\}) < \frac{1}{2}$ for $n \geqslant n_1$. Let $n_2 > n_1$ such that $\mu(\{|f_n - f| > \frac{1}{2}\}) < \frac{1}{4}$. Repeat this process, choosing $n_k > n_{k-1}$ such that $\mu(\{|f_{n_k} - f| > \frac{1}{k}\}) < \frac{1}{2^k}$.

> **Claim** $f_{n_k} \to f$ a.e.
>
> **Proof** Let $A_k = \{|f_{n_k} - f| > \frac{1}{k}\}$, $A = \{x \mid x \in A_k$ only finitely many times$\}$. By definition, if $x \in A$, then $f_{n_k}(x) \to f(x)$. Now it suffices to show that $\mu(A^c) = 0$. Write
>
> $$A^c = \{x \mid x \in A_k \text{ infinitely often}\} = \bigcap_{k\geqslant 1} \bigcup_{n\geqslant k} A_n$$

Now for any $k$,

$$\mu(A^c) \leqslant \mu(\bigcup_{n\geqslant k} A_n) \leqslant \sum_{n\geqslant k} \tfrac{1}{2^n} = \tfrac{1}{2^{k-1}}$$

Let $k \to \infty$, then $\mu(A^c) = 0$. □

□

Now let's investigate $L^p$ convergence.

Definition - Let $V$ a vector space. We say $\|\cdot\|$ is a *norm* on $V$ if

1. $\|v\| \geqslant 0$ for all $v \in V$ and $\|v\| = 0$ if and only if $v = 0$.
2. For all $\alpha \in \mathbb{R}$, $\|\alpha v\| = |\alpha|\|v\|$.
3. $\|u + v\| \leqslant \|u\| + \|v\|$ for all $u, v \in V$ (triangle inequality).

Given a norm $\|\cdot\|$ on $V$, define $d : V \times V \to \mathbb{R}$ by $d(u, v) = \|u - v\|$. It's not hard to check that $d$ is a metric, so now $(V, d)$ is a metric space.

Definition - Define $\mathcal{L}^1(X, \Sigma, \mu) = \{f : X \to [-\infty, \infty] \mid \int_X |f|\, d\mu < \infty\}$. Clearly $\mathcal{L}^1$ is a vector space. We attempt to define a norm on $\mathcal{L}^1$ by

$$\|f\|_1 = \int_X |f|\, d\mu$$

This clearly satsfies properties (2) and (3) of norms, but $\|f\|_1 = 0$ only implies that $f = 0$ a.e., not that $f = 0$. This motivates the following definition.
Definition - Define an equivalence relation $\sim$ on $\mathcal{L}^1$ by $f \sim g$ if $f = g$ a.e. Set $L^1 = \mathcal{L}^1/\sim$. Now for $F \in L^1$, $F$ represented by some function $f$, define $\|F\|_1 = \|f\|_1$. This is well-defined by definition of $\sim$.

**Lecture 16 - 10/01/2014**

Recall - In $L^p$ spaces, convergence $f_n \xrightarrow{L^p} f$ is defined as $\lim_{n\to\infty} \int_X |f_n - f|^p \, d\mu = 0$.
Exercise - $|f| \leqslant \|f\|_\infty$ a.e.

Henceforth, assume that $(X, \Sigma, \mu)$ is complete.
Definition - $\mathcal{L}^p(X,\mu) := \{f : X \to [-\infty,\infty] \mid \|f\|_p < \infty\}$.

Note that $\|f\|_p = 0$ only implies $f = 0$ a.e., not $f = 0$. To make $\|\cdot\|_p$ a norm, define an equivalence relation $\sim$ on $\mathcal{L}^p$ by $f \sim g$ if and only if $f = g$ a.e. Now define $L^p(X,\mu) = \mathcal{L}^p(X,\mu)/\sim$. For $F \in \mathcal{L}^p$, define $\|F\|_p = \|f\|_p$, where $f$ is any representative of $F$. This is well-defined by definition of $\sim$.

Given $F, G \in L^p(X,\mu)$, define $F + G = [f+g]$, where $F = f/\sim$, $G = g/\sim$. This is well-defined.
Remark - We can't meaningfully talk about $F(x)$ since $f \sim g$ doesn't imply $f(x) = g(x)$. We *can* meaningfully define $\int_A F$, $\{F > a\}$, etc. In general, we treat elements of $L^p$ as functions, only doing operations that are constant on equivalence classes.

Goal 1 - $L^p$ is a Banach space.
Definition - Let $(V, \|\cdot\|)$ be a vector space, $d(u,v) = \|u - v\|$. Now $(V, d)$ is a metric space. We say $(V, \|\cdot\|)$ is a *Banach space* if $(V, d)$ is complete.
E.g. $\mathbb{R}^d$ is a Banach space (with the standard norm).

Theorem For all $p \in [1,\infty]$, $L^p(X,\mu)$ is a Banach space.
Proof - We only need to show

1. $\|f+g\|_p \leqslant \|f\|_p + \|g\|_p$
2. Completeness

The rest is easy. IOU this proof. □

Lemma (Hölder's Inequality) - Let $p, q \in [1,\infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^p$, $g \in L^q$, then $|\int fg| \leqslant \|f\|_p \|g\|_q$. In particular, $fg \in L^1$.
Remark - $p = q = 2$ is Cauchy-Schwarz.
Intuition - The relationship between $p$ and $q$ follows by "counting dimensions". Say $X = \mathbb{R}^d$, $f, g$ are dimensionless quantities. Then $\int_{\mathbb{R}^d} fg$ has dimension $L^d$, $\|f\|_p$ has dimension $L^{d/p}$, and $\|g\|_q$ has dimension $L^{d/q}$. We want $L^d = L^{d/p}L^{d/q}$, which holds if and only if $\frac{1}{p} + \frac{1}{q} = 1$.

To formalize this argument (see homework), replace $f(x)$ by $f(\lambda x)$. Do a change of variables and then take $\lambda$ large and $\lambda$ small to see that the inequality only holds when $\frac{1}{p} + \frac{1}{q} = 1$.

Induction Proof (Hölder) - The main idea is to show that for real numbers $x_i, y_i$, $i = 1, \ldots, N$, we have $|\sum_i x_i y_i| \leqslant (\sum_i |x_i|^p)^{1/p} (\sum_i |y_i|^q)^{1/q}$. Now let $s, t$ be simple functions and write $s = \sum_i x_i \chi_{E_i}$, $t = \sum_i y_i \chi_{E_i}$, where $x_i, y_i \in \mathbb{R}$ and the sets $\{E_i\}$ are disjoint. Writing $\mu(E_i) = \mu(E_i)^{1/p}\mu(E_i)^{1/q}$, we have

$$\|st\|_1 = |\sum_i \mu(E_i) x_i y_i| \leqslant |\sum_i |x_i|^p \mu(E_i)|^{1/p} |\sum_i |y_i|^q \mu(E_i)|^{1/q} = \|s\|_p \|t\|_q$$

Now the result holds for simple functions, so use approximation and monotone convergence to finish this off. □

Better Proof - First assume that $p, q \in (1,\infty)$. The proof starts with Young's Inequality, which is as follows.

**Claim (Young's Inequality)** Let $x, y \geqslant 0$ and let $p, q \in (1, \infty)$, $\frac{1}{p} + \frac{1}{q} = 1$. Then $xy \leqslant \frac{x^p}{p} + \frac{y^q}{q}$.

**Proof** This follows from convexity of the function $h(z) = \ln(z)$, since

$$\ln(xy) = \frac{\ln(x^p)}{p} + \frac{\ln(x^q)}{q} \leqslant \ln(\frac{x^p}{p} + \frac{y^q}{q})$$

□

To prove Hölder, first note that if $\|f\|_p = 0$ or $\|g\|_q = 0$, then $f = 0$ or $g = 0$, so we are done. Assume now that $\|f\|_p, \|g\|_q \neq 0$. Set $F = \frac{f}{\|f\|_p}$, $G = \frac{g}{\|g\|_q}$, so that $\|F\|_p = \|G\|_q = 1$. By monotonicity and Young's inequality, we have

$$|\int FG| \leqslant \int |F||G| \leqslant \frac{\int |F|^p}{p} + \frac{\int |G|^q}{q} = 1$$

Multiplying by $\|f\|_p \|g\|_q$, we have $|\int fg| \leqslant \|f\|_p \|g\|_q$.

The case $p = 1, q = \infty$ is left as an exercise. □

Lemma (Duality) - Let $p \in [1, \infty)$, $f \in L^p$, $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$\|f\|_p = \sup_{g \in L^q \setminus \{0\}} \frac{1}{\|g\|_q} \int fg \quad (*)$$

Remark - If $p = \infty$, this is still true provided that $X$ is $\sigma$-finite.
Proof - By Hölder, for any $g \in L^q \setminus \{0\}$, we have $\frac{1}{\|g\|_q} \int fg \leqslant \|f\|_p$. Thus, LHS $\geqslant$ RHS in $(*)$. For the reverse inequality, set $F = \frac{f}{\|f\|_p}$. It's enough to show that $\sup_{\|G\|_q = 1} \int FG = 1$. To get this, choose $G = |F|^{p-1}\text{sign}(F)$. Clearly $\int FG = \int |F|^p = 1$. Moreover,

$$\int |G|^q = \int |F|^{(p-1)q} = \int |F|^{pq-q} = \int |F|^p = 1$$

□

**Lecture 17 - 10/03/2014**

IOU - $L^p$ is a Banach space. We still need the triangle inequality and completeness. We'll prove the first today.

Last time, we saw the following:

1. If $p, q \in [1, \infty]$, $\frac{1}{p} + \frac{1}{q} = 1$, then $|\int fg| \leqslant \|f\|_p \|g\|_q$.
2. We also have a duality principle: Given $p \in [1, \infty)$, $f \in L^p$, we have
$$\|f\|_p = \sup_{q \in L^q \setminus \{0\}} \frac{1}{\|g\|_q} \int fg$$
3. The previous statement holds for $p = \infty$ if $X$ is $\sigma$-finite.

Proposition - Let $f, g \in L^p$. Then $f + g \in L^p$ and

$$\|f + g\|_p \leqslant \|f\|_p + \|g\|_p$$

Proof - If $p = \infty$, this is easy. So assume $p \in [1, \infty)$. The function $x \mapsto |x|^p$ is convex, so

$$\frac{1}{2^p}|f + g|^p = |\tfrac{1}{2}f + \tfrac{1}{2}g|^p \leqslant \tfrac{1}{2}|f|^p + \tfrac{1}{2}|g|^p$$

Since the right-hand side is integrable, so is the left-hand side, thus $f + g \in L^p$.

For the triangle inequality, let $q$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. By duality, we have

$$\begin{aligned}
\|f + g\|_p &= \sup_{h \in L^q \setminus \{0\}} \frac{1}{\|h\|_q} \int (f + g)h \\
&\leqslant \sup_{h \in L^q \setminus \{0\}} \frac{1}{\|h\|_q} \int fh + \sup_{h \in L^q \setminus \{0\}} \frac{1}{\|h\|_q} \int gh \\
&= \|f\|_p + \|g\|_p
\end{aligned}$$

□

Quote - "Duality gets you an amazing amount of mileage."

Before we prove completeness, let's discuss another useful inequality. On your homework, you'll show that this can be used to prove Hölder's inequality. First recall that if $\phi : \mathbb{R} \to \mathbb{R}$ is convex, then given a finite set of points $c_i \in [0, 1]$ with $\sum_i c_i = 1$, we have $\phi(\sum_i c_i x_i) \leqslant \sum_i c_i \phi(x_i)$. Loosely speaking, Jensen's inequality extends this result to integrals.

Jensen's Inequality - Let $\phi$ be a convex real-valued function, $\mu(X) = 1$. Then for any $f \in L^1(X)$, we have
$$\phi(\int_X f \, d\mu) \leqslant \int_X \phi \circ f \, d\mu$$

Remark - If $\phi$ is convex, then $\phi$ is continuous. In particular, $\phi$ is Borel, so $\phi \circ f$ is measurable. (This justifies well-definedness of the expression on the right.)

Completeness of $L^p$

Lemma - Let $p \in [1, \infty)$. Say $(f_n) \subseteq L^p$ such that $\sum_n \|f_n\|_p < \infty$. Then

1. $\sum_n f_n$ converges pointwise a.e., say to $f$.

2. $f \in L^p$ and $\sum\limits_n f_n$ converges in $L^p$ to $f$.

Proof - Define $s_N = \sum\limits_{1 \leqslant n \leqslant N} f_n$, $t_N = \sum\limits_{1 \leqslant n \leqslant N} |f_n|$. For (1), we must show that $s_N$ converges and is finite a.e., for which it's enough to show that $t_N$ converges and is finite a.e.

Since the functions $(t_n)$ are increasing and nonnegative, we can set $F = \sum\limits_n |f_n| = \lim\limits_{N\to\infty} t_N$. By the triangle inequality, we have

$$\|t_N\|_p \leqslant \sum_{1 \leqslant n \leqslant N} \|f_n\|_p \leqslant \sum_n \|f_n\|_p < \infty$$

Now $\int t_N^p = \|t_N\|_p^p \leqslant (\sum\limits_n \|f_n\|_p)^p < \infty$, so by the MCT, we have

$$\lim_{N\to\infty} \int t_N^p = \int F^p < \infty$$

Thus $F \in L^p$. In particular, $F$ is finite a.e., so now (1) holds.

For (2), note that $|f| \leqslant F$ implies $\|f\|_p \leqslant \|F\|_p$, so now $f \in L^p$. Moreover, by the MCT and the triangle inequalities for $L^p$ and $\mathbb{R}$, we have

$$\begin{aligned} \|f - s_N\|_p = \|\sum_{n \geqslant N+1} f_n\|_p \leqslant \|\sum_{n \geqslant N+1} |f_n|\|_p &= \lim_{m\to\infty} \|\sum_{m \geqslant n \geqslant N+1} |f_n|\|_p \\ &\leqslant \lim_{m\to\infty} \sum_{m \geqslant n \geqslant N+1} \|f_n\|_p \\ &= \sum_{n \geqslant N+1} \|f_n\|_p \to 0 \text{ as } N \to \infty \end{aligned}$$

Thus, $s_N \to f$ in $L^p$. □

Proposition - $L^p$ is complete.
Proof - Let $(f_n) \subseteq L^p$ be Cauchy. It's enough to find a subsequence $(f_{n_k})$ of $(f_n)$ such that $(f_{n_k})$ converges in $L^p$. To do this, choose $n_1 < n_2 < \cdots$ such that $\|f_{n_{k+1}} - f_{n_k}\|_p < \frac{1}{2^k}$. By the lemma, since $\sum\limits_k \|f_{n_{k+1}} - f_{n_k}\|_p < 1$, then $\sum\limits_k (f_{n_{k+1}} - f_{n_k})$ converges both a.e. and in $L^p$. Set

$$f = \sum_k (f_{n_{k+1}} - f_{n_k}) + f_{n_1}$$

Let the sum telescope, so that $f_{n_k} \xrightarrow{L^p} f$. □
Exercise - Prove that $L^\infty$ is complete.

Convergence - We have convergence a.e., convergence in measure, convergence in $L^p$. What's the relationship between these notions?

- We know a.e. convergence implies convergence in measure if $\mu(X) < \infty$ (but not in general).
- We know convergence in measure implies a.e. convergence along a subsequence.
- In general, a.e. convergence does not imply convergence in $L^p$, even on finite measure spaces. For example, $X = (0, 1)$, $f_n = 2^n \chi_{(0,\frac{1}{n})}$.
- However, convergence in $L^p$ does imply a.e. convergence along a subsequence (similar to second item above).

- Convergence in $L^p$ implies convergence in measure (next time, using Chebyshev).
- Convergence in measure does not imply convergence in $L^p$ (next time).

Lemma (Chebyshev's Inequality) - Let $f$ be meaurable. Then for all $\lambda > 0$,

$$\mu(|f| > \lambda) \leqslant \frac{1}{\lambda}\|f\|_1$$

Proof - Clearly $\lambda\chi_{|f|>\lambda} \leqslant |f|$, so integrate both sides and we're done. $\square$

**Lecture 18 - 10/06/2014**

Proposition - $(f_n) \xrightarrow{L^p} f$, $p \in [1, \infty)$ implies $f_n \xrightarrow{\mu} f$.
Proof - Use Chebyshev. We have

$$\mu(\{|f_n - f| > \epsilon\}) = \mu(\{|f_n - f|^p > \epsilon^p\}) \leqslant \tfrac{1}{\epsilon^p}\|f_n - f\|_p^p \to 0 \text{ as } n \to \infty$$

IOU - $f_n \xrightarrow{\mu} f$ + "mystery assumption" is equivalent to $f_n \xrightarrow{L^p} f$.

Proposition 1 - $\mathcal{S} = \{s \mid s \text{ simple}, \ \mu(\{s \neq 0\}) < \infty\}$ is dense in $L^p$ for all $p \in [1, \infty]$.
Proof - Let $f \in L^p$. If $p \in [1, \infty)$, let $(s_n)$ be sequence of simple functions with $s_n \to f$ a.e. and $|s_n| \leqslant f$ a.e. Note, $f \in L^p$ implies $s_n \in L^p$, so then $\mu(\{s_n \neq 0\}) < \infty$ for each $n$. Now by the LDCT,

$$\|s_n - f\|_p^p = \int |s_n - f|^p \to 0$$

since $|s_n - f|^p \leqslant (|f| + |f|)^p = 2^p|f|^p$, which is integrable.

If $p = \infty$, then partition $[-\|f\|_\infty, \|f\|_\infty]$ into finitely many subintervals $\{A_i\}$ of length at most $\epsilon$, then set $s = \sum_i x_i \chi_{f^{-1}(A_i)}$, where $x_i$ is the midpoint of $A_i$. Now $|f - s| < \epsilon$ a.e., so $\|f - s\|_\infty \leqslant \epsilon$. □

Proposition 2 - Let $X$ be a $\sigma$-compact metric space (i.e. $X = \bigcup_n K_n$, where the sets $(K_n)$ are increasing and compact). If $\mu$ is regular on $X$, then $C_c(X)$ is dense in $L^p$.
Remark/Definition - $C_c(X) = \{$continuous functions on $X$ with compact support$\}$.
Proof - First assume that $X$ is compact. Let $f \in L^p$ and fix $\epsilon > 0$. For each $n \geqslant 1$, define

$$f_n = \begin{cases} f & |f| \leqslant n \\ n & f > n \\ -n & f < -n \end{cases}$$

(Think of this as a truncated version of $f$.) By LDCT, $f_n \xrightarrow{L^p} f$ since $f_n \leqslant |f|$. Let $N$ sufficiently large so that $\|f_N - f\|_p < \frac{\epsilon}{2}$. Choose $\delta > 0$ sufficiently small so that $\delta(2n)^p < (\frac{\epsilon}{2})^p$. By Lusin's theorem, there is continuous $g_n$ such that $\mu(\{f_n \neq g_n\}) < \delta$. By truncating if necessary, we may assume $|g_n| \leqslant n$, so now $g_n$ has compact support. Moreover,

$$\|f_n - g_n\|_p^p = \int_{\{f_n \neq g_n\}} |f_n - g_n|^p \leqslant \delta(2n)^p < (\tfrac{\epsilon}{2})^p$$

Now $\|g_n - f\|_p \leqslant \|g_n - f_n\|_p + \|f_n - f\|_p < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$, so we're done.

The case that $X$ is $\sigma$-compact is left as an exercise. □

Proposition 3 - If $\Sigma = \sigma(\mathcal{C})$, where $\mathcal{C}$ is countable, then $L^p(X)$ is separable for $p \in [1, \infty)$.
Proof - Homework. □

Theorem (Vitali) - Let $f \in L^1(X)$, $(f_n) \subseteq L^1(X)$. Then $f_n \xrightarrow{L^1} f$ if and only if

1. $f_n \xrightarrow{\mu} f$
2. For any $\epsilon > 0$, there is $\delta > 0$ such that $\mu(A) < \delta$ implies $\int_A |f_n| < \epsilon$ for all $n$.
3. For any $\epsilon > 0$, there is $F \in \Sigma$ such that $\mu(F) < \infty$ and $\int_{F^c} |f_n| < \epsilon$ for all $n$.

Quotes "Condition (2) prevents mass collecting at a point. Condition (3) prevents mass escaping to infinity."
Remark - Condition (2) is sometimes referred to as *equi-integrability*. Condition (3) is sometimes referred to as *tightness*.
Proof (of Vitali) - Suppose that $f_n \to f$ in $L^1$. We already have $f_n \xrightarrow{\mu} f$.

To show equi-integrability, let $n_0$ be such that $n > n_0$ implies $\int_X |f - f_n| < \frac{\epsilon}{2}$. Now for any set $A \subseteq X$, we have

$$\int_A |f_n| \leqslant \int_A |f - f_n| + \int_A |f| < \int_A |f| + \tfrac{\epsilon}{2}$$

If $\mu(A) < \delta$, then for any $\lambda > 0$,

$$\int_A |f| = \int_{A\cap\{|f|\leqslant\lambda\}} |f| + \int_{A\cap\{|f|>\lambda\}} |f| \leqslant \lambda\delta + \int_{\{|f|>\lambda\}} |f|$$

By LDCT, we have $\lim\limits_{\lambda\to\infty} \int_{\{|f|>\lambda\}} |f| = \lim\limits_{\lambda\to\infty} \int_X \chi_{\{|f|>\lambda\}}|f| = 0$. Now let $\lambda$ be sufficiently large such that $\int_{\{|f|>\lambda\}} |f| < \frac{\epsilon}{4}$ and let $\delta > 0$ be sufficiently small such that $\lambda\delta < \frac{\epsilon}{4}$. Now

$$\mu(A) < \delta,\ n > n_0 \implies \int_A |f_n| < \epsilon$$

To get the result for $n \leqslant n_0$, repeat the above argument (the one used to minimize $\int_A |f|$) for the finitely many functions $f_n$, $n \leqslant n_0$.

Next time, we'll prove that $(f_n)$ is tight and prove the converse of the theorem. □

**Lecture 19 - 10/08/2014**

Proof (of Vitali, continued) - We need to show that $f_n \xrightarrow{L^1} f$ implies $(f_n)$ is tight. Fix $\epsilon > 0$. For any $F$, we have

$$\int_F |f_n| \leqslant \int_F |f| + \int_F |f_n - f| < \int_F |f| + \tfrac{\epsilon}{2}$$

for $n$ sufficiently large, say $n \geqslant n_0$ (where $n_0$ is independent of $F$).

Set $F = \{|f| > \lambda\}$ (we'll choose $\lambda$ later). By Chebyshev's inequality, we have $\mu(F) \leqslant \frac{\|f\|_1}{\lambda} < \infty$. Now $\int_{F^c} |f| = \int |f|\chi_{\{|f|\leqslant\lambda\}}$. As $\lambda \to 0^+$, $\int |f|\chi_{\{|f|\leqslant\lambda\}} \to 0$ by LDCT, so let $\lambda$ be sufficiently small so that $\int |f|\chi_{\{|f|\leqslant\lambda\}} < \frac{\epsilon}{2}$. Repeat the same trick for the finitely many functions $f_1, \ldots, f_{n_0}$, then we're done.

Conversely, suppose that $f_n \xrightarrow{\mu} f$, $(f_n)$ is equi-integrable, and $(f_n)$ is tight. Fix $\epsilon > 0$. Then for any set $F$ and $\lambda > 0$, we have

$$\begin{aligned}\int |f_n - f| &= \int_F |f_n - f| + \int_{F^c} |f_n - f| \\ &= \int\limits_{F\cap\{|f_n - f|\leqslant\lambda\}} |f_n - f| + \int\limits_{F\cap\{|f_n-f|>\lambda\}} |f_n - f| + \int_{F^c} |f_n - f| \quad (*)\end{aligned}$$

Since $(f_n)$ is tight and $f \in L^1$, then $(f_n) \cup \{f\}$ is tight. Choose $F$ such that $\mu(F) < \infty$ and $\int_{F^c}(|f| + |f_n|) < \frac{\epsilon}{3}$, which is an upper bound for the third term of $(*)$.

The first term is bounded above by $\lambda\mu(F)$, so let $\lambda > 0$ sufficiently small such that $\lambda\mu(F) < \frac{\epsilon}{3}$. This bounds the first term.

Since $(f_n)$ is equi-integrable and $f \in L^1$, then $(f_n) \cup \{f\}$ is equi-integrable. Let $\delta > 0$ such that $\mu(A) < \delta$ implies $\int_A (|f_n| + |f|) < \frac{\epsilon}{3}$. Finally, since $f_n \xrightarrow{\mu} f$, let $n_0$ be sufficiently large such that $n > n_0$ implies $\mu(\{|f_n - f| > \lambda\}) < \delta$. This bounds the second term above by $\frac{\epsilon}{3}$, finishing the proof. ∎

Remark - If $\mu(X) < \infty$, then any family of $L^1$ functions is tight.

Remark - Let $F \in L^1$, $|f_n| \leqslant F$ for all $n$. Then

1. $(f_n)$ is equi-integrable.
2. $(f_n)$ is tight.

Proposition

1. Say $\lim\limits_{\lambda\to\infty} \sup\limits_n \int_{\{|f_n|>\lambda\}} |f_n| = 0$. Then $(f_n)$ is equi-integrable.
2. (de la Vallée-Poussin) Suppose there is increasing $\phi : [0, \infty) \to [0, \infty)$ such that $\lim\limits_{x\to\infty} \frac{\phi(x)}{x} = \infty$ and $\sup\limits_n \int_X \phi(|f_n|) < \infty$. Then $(f_n)$ is equi-integrable.

Proof - (1) Fix $\epsilon > 0$. Suppose $\mu(A) < \delta$ (we'll choose $\delta > 0$ later). Then

$$\int_A |f_n| = \int\limits_{A\cap\{|f_n|\leqslant\lambda\}} |f_n| + \int_{A\cap\{|f_n|>\lambda\}} |f_n| \leqslant \lambda\mu(A) + \int\limits_{\{|f_n|>\lambda\}} |f_n|$$

Let $\lambda$ be sufficiently large so that $\int_{\{|f_n|>\lambda\}} |f_n| < \frac{\epsilon}{2}$. Now let $\delta > 0$ be sufficiently small so that $\lambda\delta < \frac{\epsilon}{2}$.

(2) By (1), it's enough to show that $\lim_{\lambda\to 0}\sup_n \int_{\{|f_n|>\lambda\}} |f_n| = 0$. Let $C = \sup_n \int_X \phi(|f_n|) < \infty$. Let $\lambda_0$ be such that $\lambda > \lambda_0$ implies $\lambda < \epsilon\phi(\lambda)$. Now

$$\int_{\{|f_n|>\lambda_0\}} |f_n| \leqslant \int_{\{|f_n|>\lambda_0\}} \epsilon\phi(|f_n|) \leqslant \epsilon C$$

Take the supremum over $n$ and then take the limit $\lambda_0 \to \infty$, so now

$$\lim_{\lambda_0\to\infty} \sup_n \int_{\{|f_n|>\lambda_0\}} |f_n| = 0$$

□

"Classic" Example - Say $\mu(X) < \infty$, $f_n \to f$ a.e., and $\sup_n \|f_n\|_p < \infty$ for some $p > 1$. Then $f_n \to f$ in $L^1$. This follows from (2) above with $\phi(x) = x^p$. Why?

Since $\mu(X) < \infty$, then a.e. convergence implies convergence in measure, and we also have tightness for free. We get equi-integrability by (2) with $\phi(x) = x^p$.

**Lecture 20 - 10/13/2014**

Plan - Signed measures, Hahn decomposition
Setting - $(X, \Sigma)$ a measurable space

Definition - Let $\mu : \Sigma \to (-\infty, \infty]$ or $[-\infty, \infty)$. We say $\mu$ is a *signed measure* if $\mu$ is countably additive and $\mu(\varnothing) = 0$.
E.g. Let $\mu_1, \mu_2$ be positive measures, $\mu_2(X) < \infty$. Then $\mu := \mu_1 - \mu_2$ is a signed measure.
E.g. 2 - Let $\nu$ be a positive measure, $f : X \to \mathbb{R}$ measurable, and suppose that $f^- = \max(0, -f) \in L^1$. Then define

$$\mu(A) = \int_A f^+ \, d\nu - \int_A f^- \, d\nu, \quad A \in \Sigma$$

$\mu$ is a signed measure.
Proof - Countable additivity follows from monotone convergence. □

Motivation - We'd like to form a normed vector space of measures. We can't do this with positive measures alone.
Remark - This construction has applications to finding solutions of stochastic differential equations.

Definition - Let $\mu$ be a signed measure on $(X, \Sigma)$. We say $A \in \Sigma$ is $\mu$-*positive* if for every $E \subseteq A$ with $E \in \Sigma$, we have $\mu(A) \geqslant 0$, i.e. when $\mu$ is restricted to $A$, we get a positive measure. We define $\mu$-negative sets similarly.
Notation - Denote $\mathcal{P}(\mu) = \{\mu\text{-positive sets}\}$, $\mathcal{N}(\mu) = \{\mu\textit{-negative sets}\}$.

Remarks

1. If $A \in \Sigma$ such that $\mu(A)$ is finite, then $\mu(E)$ is finite for any $E \subseteq A$ in $\Sigma$. This follows from writing
$$\mu(A) = \mu(E) + \mu(A \backslash E)$$
2. $\mathcal{P}(\mu)$ and $\mathcal{N}(\mu)$ are closed under countable unions and intersections, as well as (relative) complements.
3. For any $P \in \mathcal{P}(\mu)$ and $N \in \mathcal{N}(\mu)$, $P$ and $N$ are $\mu$-disjoint. That is, for all $E \subseteq P \cap N$, $\mu(E) = 0$.

Theorem (Hahn Decomposition) - Let $\mu$ be a signed measure on $(X, \Sigma)$. Then there exist $P \in \mathcal{P}(\mu)$ and $N \in \mathcal{N}(\mu)$ such that $X = P \cup N$.
Proof - Without loss of generality, assume $\mu > -\infty$. Denote

$$L = \inf\{\mu(A) : A \in \mathcal{N}(\mu)\}$$

(Note: There is at least one such $A$, namely $A = \varnothing$.)

**Claim** $L$ is finite. Moreover, the infimum is witnessed.

**Proof** Let $(A_n) \subseteq \mathcal{N}(\mu)$ such that $\mu(A_n) \to L$. Consider the union $N = \bigcup_n N_n$. By a previous remark, $N \in \mathcal{N}(\mu)$. Thus, $L \leqslant \mu(N)$ by definition. On the other hand, $\mu(N) = \mu(A_n) + \mu(N \backslash A_n) \leqslant \mu(A_n)$ for each $n$, so $\mu(N) \leqslant L$. Thus, $\mu(N) = L$. In particular, $L$ is finite (since $\mu > -\infty$).

Set $P = X\backslash N$. It remains to show that $P \in \mathcal{P}(\mu)$. We show this by contradiction.

Suppose that $P \notin \mathcal{P}(\mu)$. Let $A \subseteq P$ in $\Sigma$ such that $\mu(A) < 0$. If $A \in \mathcal{N}(\mu)$, we have a contradiction, since

$$\mu(A \cup N) \leqslant \mu(A) + \mu(N) < \mu(N) = L,$$

contradicting the definition of $L$. If not, the following lemma will yield the same contradiction.

**Lemma** Let $A \in \Sigma$, $-\infty < \mu(A) < 0$. Then there is $E \subseteq A$, $E \in \mathcal{N}(\mu)$, such that $\mu(E) \leqslant \mu(A)$.

**Proof** The idea is to cut out the parts of $A$ that are positive. We do this as follows. Set

$$\delta_1 = \sup\{\mu(E) : E \subseteq A,\ E \in \Sigma\}$$

If $\delta_1 = 0$, then $A \in \mathcal{N}(\mu)$, so take $E = A$. If not, $\delta_1 \in (0, \infty]$, so choose $A_1 \subseteq A$ in $\Sigma$ such that $\mu(A_1) \geqslant \min(\frac{\delta_1}{2}, 1)$.

Continue recursively. At step $n$, we have

(a) $\delta_1 \geqslant \delta_2 \geqslant \cdots \geqslant \delta_{n-1} > 0$

(b) disjoint subsets $(A_k)_{k=1}^{n-1}$ of $A$ such that $\mu(A_k) \geqslant \min(\frac{\delta_k}{2}, 1)$.

Define $\delta_n = \sup\{\mu(E) : E \subseteq A\backslash(\bigcup_i A_i), E \in \Sigma\}$. If $\delta_n = 0$, stop and set $E = A\backslash(\bigcup_i A_i)$. Otherwise, $\delta_n \in (0, \infty]$, so let $A_n \subseteq (A\backslash \bigcup_i A_i)$ with $\mu(A_n) \geqslant \min(\frac{\delta_n}{2}, 1)$.

If this process stops at any point, we are done. Otherwise, set $E := A\backslash \bigcup\limits_{n\geqslant 1} A_n$. Why does this work? First write

$$\mu(E) = \mu(A) - \sum_n \mu(A_n) < \mu(A) \qquad (*)$$

It remains to argue that $E \in \mathcal{N}(\mu)$. This boils down to showing that $\delta_n \to 0$ as $n \to \infty$. Let $D \subseteq E$ in $\Sigma$ with $\mu(D) \geqslant 0$. Then for each $n$, $D \subseteq A\backslash \bigcup\limits_{1\leqslant k\leqslant n} A_k$. In particular, $\mu(D) \leqslant \delta_n$. We are done, provided we show that $\delta_n \to 0$. But this follows from $(*)$ above. Since $\mu(E)$ is finite, then $\sum\limits_n \mu(A_n)$ is finite, so $\sum\limits_n \min(\frac{\delta_n}{2}, 1) < \infty$, so $\delta_n \to 0$. □

□

**Lecture 21 - 10/15/2014**

Plan: Jordan decomposition, $\mathcal{M}(X, \Sigma)$
Setting - $(X, \Sigma)$ is a measurable space, $\mu$ a signed measure with Hahn decomposition $(P, N)$, where $P \cap N = \varnothing$, $X = P \cup N$, $P$ $\mu$-positive and $N$ $\mu$-negative.

Definition - We say signed measures $\mu_1$ or $\mu_2$ are *singular*, denoted $\mu_1 \perp \mu_2$, if there exist disjoint $X_1, X_2 \in \Sigma$ such that for all $A \in \Sigma$,

$$\mu_1(A) = \mu_1(A \cap X_1), \qquad \mu_2(A) = \mu_2(A \cap X_2)$$

In this case, we say $\mu_1$ is *concentrated on* $X_1$ and $\mu_2$ is *concentrated on* $X_2$.
Theorem (Jordan Decomposition) - Given $\mu$, there exist unique (positive) measures $\mu^+$ and $\mu^-$ such that

(a) $\mu = \mu^+ - \mu^-$.

(b) At least one of $\mu^+$, $\mu^-$ is finite.

(c) $\mu^+ \perp \mu^-$.

Remark This uniqueness is in some sense more advantageous than the Hahn decomposition, since we may able to change the sets $P$ and $N$ in the Hahn decomposition, i.e. $P$ and $N$ are only *essentially* unique. By contrast, $\mu^+$ and $\mu^-$ are unique by (c).
Proof - Existence is easy. Let $(P, N)$ be a Hahn decomposition for $\mu$. Define

$$\mu^+(A) = \mu(A \cap P), \qquad \mu^-(A) = -\mu(A \cap N)$$

By definition, $\mu^+$ and $\mu^-$ satisfy (a)-(c).
For uniqueness, let $\mu^+$ and $\mu^-$ satisfy (a)-(c). We show that for any $A \in \Sigma$, we have

$$\mu^+(A) = \sup_{E \subseteq A, E \in \Sigma} \mu(E) \qquad (*)$$

($\geqslant$) For any $E \subseteq A$ in $\Sigma$, we have

$$\mu^+(A) \geqslant \mu^+(E) = \mu(E) + \mu^-(E) \geqslant \mu(E)$$

($\leqslant$) Let $(X^+, X^-)$ be such that $\mu^+$ is concentrated on $X^+$ and $\mu^-$ is concentrated on $X^-$. Now

$$\mu^+(A) = \mu^+(A \cap X^+) = \mu(A \cap X^+) + \mu^-(A \cap X^+) = \mu(A \cap X^+)$$

Take $E = A \cap X^+$ and $(*)$ is proven.
From $(*)$, we get $\mu^+$. To get $\mu^-$, simply apply $(*)$ again, replacing $\mu$ by $-\mu$. This works because $-\mu = \mu^- - \mu^+$, so $(-\mu)^+ = \mu^-$. □

Definition - Given the Jordan decomposition $\mu = \mu^+ - \mu^-$, the *total variation* of $\mu$ is defined by

$$|\mu| := \mu^+ + \mu^-$$

If $|\mu|(X) < \infty$, then $\|\mu\| = |\mu|(X)$ is called the *total variation norm* of $\mu$.
Remark - We'll se that $\|\cdot\|$ is a norm on the space of bounded total variation measures.
E.g. - Let $\nu$ be a measure, $f \in L^1(\nu)$. Define $\mu(A) = \int_A f\, d\nu = \int_A f^+\, d\nu - \int_A f^-\, d\nu$. In this case,

$$|\mu|(A) = \int_A |f|\, d\nu, \qquad \|\mu\| = \int_X |f|\, d\nu = \|f\|_{L^1(\nu)}$$

Remark/Quote - "The space of signed measures is an *extension* of $L^1$-space."

Theorem - The space $\mathcal{M}(X,\Sigma)$ of finite signed measures, equipped with the total variation norm, is a Banach space.
Proof - Everything but Banach is easy. Let $(\mu_n) \subseteq \mathcal{M}(X,\Sigma)$ be Cauchy, i.e.

$$\|\mu_n - \mu_m\| \to 0 \quad \text{as } n,m \to \infty \qquad (**)$$

In this case, $\sup_n \|\mu_n\| < \infty$ (by the triangle inequality, Cauchy sequences are bounded).

Observe that for any $\nu \in \mathcal{M}(X,\Sigma)$, we have

$$\begin{aligned}\|\nu\| = \|\nu^+\| + \|\nu^-\| &= \nu^+(X) - \nu^-(X) \\ &= \sup_{\substack{\text{disjoint} \\ A_1,A_2 \in \Sigma}} \{\nu(A) - \nu(A_2)\} \\ &= \sup_{\substack{\text{disjoint} \\ A_1,A_2 \in \Sigma}} \{|\nu(A_1)| + |\nu(A_2)|\} \\ &\leqslant 2\sup_{A\in\Sigma} |\nu(A)|\end{aligned}$$

In particular, we have shown that $\sup_{A\in\Sigma} |\nu(A)| \leqslant \|\nu\| \leqslant 2\sup_{A\in\Sigma} |\nu(A)|$. Thus, $(**)$ is equivalent to

$$\sup_{A\in\Sigma} |\mu_n - \mu_m(A)| \to 0 \quad \text{as } n,m \to \infty \qquad (***)$$

Thus for fixed $A \in \Sigma$, $(\mu_n(A))$ is a Cauchy sequence of real numbers, so define $\mu(A) := \lim_{n\to\infty} \mu_n(A)$. By $(***)$, we have

$$\sup_{A\in\Sigma} |\mu(A) - \mu_n(A)| \to 0 \text{ as } n,m \to \infty$$

It remains to show that $\mu$ is a signed measure.

To get finite additivity, it suffices to consider two disjoint sets $A, B \in \Sigma$. We have

$$\begin{aligned}&|\mu(A \cup B) - (\mu(A) + \mu(B))| \\ \leqslant\ &|\mu(A\cup B) - \mu_n(A\cup B)| + |\mu_n(A) - \mu(A)| + |\mu_n(B) - \mu(B)| \\ \leqslant\ &3\sup_{E\in\Sigma} |\mu(E) - \mu_n(E)| \to 0\end{aligned}$$

as $n \to \infty$. Now for $\sigma$-additivity, let $(A_k) \subseteq \Sigma$ be disjoint. Write

$$\begin{aligned}&|\mu(\bigcup_k A_k) - \sum_{k=1}^m \mu(A_k)| \\ \leqslant\ &(m+1)\sup_{E\in\Sigma} |\mu(E) - \mu_n(E)| + |\mu_n(\bigcup_{k\geqslant m+1} A_k)| \\ \leqslant\ &(m+1)\sup_{E\in\Sigma} |\mu(E) - \mu_n(E)| + \sup_{E\in\Sigma} |\mu_n(E) - \mu_\ell(E)| + |\mu_\ell(\bigcup_{k\geqslant m+1} A_k)|\end{aligned}$$

Take $\limsup_n$ of the right-hand side and the first term disappears. Then take $m \to \infty$ and the third term disappears. Finally, take the limit $\ell \to \infty$ so that the middle term disappears. Thus, $\mu$ is $\sigma$-additive. $\square$

**Lecture 22 - 10/20/2014**

Definition - Let $\mu, \nu$ be positive measures. We say $\nu$ is *absolutely continuous* with respect to $\mu$, denoted $\nu \ll \mu$, provided that

$$\mu(A) = 0 \implies \nu(A) = 0 \text{ for all } A \in \Sigma$$

Example - Let $(X, \Sigma, \mu)$ be a measure space, $f : X \to [0, \infty]$ measurable, and define $d\nu = f\ d\mu$, i.e. $\nu(A) = \int_A f\ d\mu$. (This was on a previous homework.)

Theorem (Radon-Nikodym) - Let $\mu, \nu \geqslant 0$ be $\sigma$-finite, $\nu \ll \mu$. Then there exists a unique (up to $\mu$-a.e.) measurable function $g : X \to [0, \infty)$ such that $d\nu = g\ d\mu$.
Notation - $g$ is called the *Radon-Nikodym derivative* of $\nu$ with respect to $\mu$. We write $g = \frac{d\nu}{d\mu}$.
Proof - **Case I:** Suppose $\mu, \nu < \infty$. Define

$$\mathcal{F} = \{f : X \to [0, \infty] \mid f \text{ measurable and } \int_A f\ d\mu \leqslant \nu(A) \text{ for all } A \in \Sigma\}$$

Note, $\mathcal{F} \neq \varnothing$ since $0 \in \mathcal{F}$.

**Claim 1** For any $f_1, f_2 \in \mathcal{F}$, $\max(f_1, f_2) \in \mathcal{F}$.

**Proof** For all $A \in \Sigma$,

$$\begin{aligned}\int_A \max(f_1, f_2) &= \int_{A \cap \{f_1 > f_2\}} f_1 + \int_{A \cap \{f_1 \leqslant f_2\}} f_2 \\ &\leqslant \nu(A \cap \{f_1 > f_2\}) + \nu(A \cap \{f_1 \leqslant f_2\}) \\ &= \nu(A)\end{aligned}$$

□

**Claim 2** Let $(f_n) \subseteq \mathcal{F}$ be increasing. Then $\lim_{n \to \infty} f_n \in \mathcal{F}$.

**Proof** Follows by Monotone Convergence Theorem. □

Now let $\alpha = \sup_{f \in \mathcal{F}} \int_X f\ d\mu \leqslant \nu(X) < \infty$. Let $(f_n) \subseteq \mathcal{F}$ with $\int_X f_n\ d\mu \to \alpha$ as $n \to \infty$. By Claim 1, we may replace each $f_n$ with $\max_{1 \leqslant i \leqslant n} \{f_i\}$ and assume that $(f_n)$ is increasing. Set $g = \lim_{n \to \infty} f_n$, now $g \in \mathcal{F}$ by Claim 2. In particular, $\int_X g\ d\mu \leqslant \nu(X) < \infty$ implies that $g < \infty$ $\mu$-a.e., so we may assume $g : X \to [0, \infty)$.

To finish Case 1, it remains to show that $d\nu = g\ d\mu$. To do this, define $d\nu_0 = d\nu - g\ d\mu$, i.e. $\nu_0(A) = \nu(A) - \int_A g\ d\mu$ for all $A \in \Sigma$. Clearly $\nu_0 \geqslant 0$ since $g \in \mathcal{F}$. We want to show that $\nu_0 \leqslant 0$, so then $\nu_0 = 0$. It's enough to prove the following claim.

**Claim** For any $\epsilon > 0$, $\nu_0 \leqslant \epsilon\mu$.

**Proof** Fix $\epsilon > 0$. Since $\nu_0 - \epsilon\mu$ is a signed measure, there is a corresponding Hahn decomposition $X = P \cup N$. To get $\nu_0 \leqslant \epsilon\mu$, we show that $\mu(P) = 0$. Then by $\nu \ll \mu$, we have $\nu(P) = 0$, so $\nu_0 - \epsilon\mu$ is negative and we're done.

To show that $\mu(P) = 0$, it's enough to show that $g + \epsilon\chi_P \in \mathcal{F}$ (by definition of $g$). Write

$$\begin{aligned}\int_A (g + \epsilon\chi_P)\ d\mu &= \int_A g + \epsilon\mu(A \cap P) \\ &= \nu(A) - \nu_0(A) + \epsilon\mu(A \cap P) \\ &= \nu(A) - \nu_0(A \cap N) - (\nu_0 - \epsilon\mu)(A \cap P) \\ &\leqslant \nu(A)\end{aligned}$$

Thus $g + \epsilon\chi_P \in \mathcal{F}$, so $\mu(P) = 0$ and we are done.

We still need to show uniqueness. Suppose that $d\nu = g_1\ d\mu = g_2\ d\mu$. Then for all $A \in \Sigma$, we have

$$\nu(A) = \int_A g_1\ d\mu = \int_A g_2\ d\mu \implies \int_A (g_1 - g_2)\ d\mu$$

Let $A = \{g_1 > g_2\}$. We get $\mu(\{g_1 > g_2\}) = 0$. By symmetry, $\mu(\{g_1 < g_2\}) = 0$. Thus, $g_1 = g_2$ $\mu$-a.e.

**Case II:** Suppose $\mu, \nu$ are $\sigma$-finite. Write $X = \bigcup_n F_n$, where $F_n \subseteq F_{n+1}$, $\mu(F_n), \nu(F_n) < \infty$. For each $n$, define $\mu_n(A) = \mu(A \cap F_n)$, $\nu_n(A) = \nu(A \cap F_n)$. Note that $\mu_n, \nu_n < \infty$ and $\nu_n \ll \mu_n$. By Case I, for each $n$ there is a unique ($\mu_n$-a.e.) $f_n$ such that $d\nu_n = f_n\ d\mu_n$. By this uniqueness, we must have $f_{n+1} \restriction_{F_n} = f_n$ for each $n$ ($\mu_n$-a.e.). Thus, $g = \lim_{n\to\infty} f_n$ is well-defined and has the desired property (by the MCT). □

**Lecture 23 - 10/22/2014**

Lebesgue Decomposition

Definition - Let $\mu$ be a positive $\sigma$-finite measure, $\nu$ a signed finite measure. We write $\nu \ll \mu$ if

$$\mu(A) = 0 \implies |\nu|(A) = 0 \text{ for all } A \in \Sigma$$

Theorem - Let $\mu$ be positive and $\sigma$-finite, $\nu$ signed and finite, $\nu \ll \mu$. Then there exists a unique $g \in L^1(\mu)$ such that $d\nu = g\, d\mu$.

Proof - By Jordan, we can write $\nu = \nu^+ - \nu^-$. Now Radon-Nikodym yields $g^+, g^-$ such that $d\nu^\pm = g^\pm\, d\mu$. Set $g = g^+ - g^-$. □

(Boring) Example - For any finite signed measure $\mu$, we have $\mu \ll |\mu|$. By Radon-Nikodym, there is $g$ such that $d\mu = g\, d|\mu|$. But we know this already! If $X$ has Hahn decomposition $X = P \cup N$ with respect to $\mu$, then $g = \chi_P - \chi_N$.

Later, we'll see a really surprising and nontrivial application of Radon-Nikodym.

Theorem (Lebesgue Decomposition) - Let $\mu, \nu$ be positive measures, $\nu$ $\sigma$-finite. Then there exist positive measures $\nu_S$ and $\nu_{AC}$ such that

1. $\nu = \nu_S + \nu_A C$
2. $\nu_{AC} \ll \mu$
3. $\nu_S \perp \mu$

Recall - $\nu_S \perp \mu$ if and only if there is $N \in \Sigma$ such that $\mu(N) = 0$, $\nu_S(N^c) = 0$.

Proof - **Case I:** Suppose $\nu$ is finite. Set

$$\mathcal{N} = \{N \in \Sigma \mid \mu(N) = 0\}, \qquad \alpha = \sup\{\nu(N) \mid N \in \mathcal{N}\} < \infty$$

Let $(N_n) \subseteq \mathcal{N}$ with $\nu(N_k) > \alpha - \frac{1}{k}$ for each $k$. Set $N = \bigcup_k N_k$. Note that $\nu(N) = \alpha$ and $\mu(N) = 0$.

Define $\nu_S(A) = \nu(A \cap N)$, $\nu_{AC}(A) = \nu(A \cap N^c)$. Clearly $\nu = \nu_S + \nu_{AC}$. Moreover, $\nu_S \perp \mu$ since $\mu(N) = 0$ and $\nu_S(N^c) = 0$. It remains to show that $\nu_{AC} \ll \mu$.

Suppose $\mu(A) = 0$. We need $\nu_{AC}(A) = 0$, i.e. $\nu(A \backslash N) = 0$. Since $A \cup N \in \mathcal{N}$, then

$$\alpha = \nu(N) \leqslant \nu(A \cup N) \leqslant \alpha$$

Since $\nu$ is finite, then $\nu(A \backslash N) = 0$, so we're done.

**Case II:** Suppose that $\nu$ is $\sigma$-finite. Write $X = \bigcup_n F_n$, where $\nu(F_n) < \infty$ and $F_n \subseteq F_{n+1}$. Set $\nu^{(n)}(A) = \nu(A \cap F_n)$, so $\nu^{(n)} < \infty$. By Case I, there exists $N_N \subseteq F_n$ and positive measures $\nu_S^{(n)}$, $\nu_{AC}^{(n)}$ such that $\nu^{(n)} = \nu_S^{(n)} + \nu_{AC}^{(n)}$ and $\mu(N_n) = 0$, $\nu_S^{(n)}(N_n^c) = 0$.

Define $N = \bigcup_n N_n$, $\nu_S(A) = \nu(A \cap N)$, $\nu_{AC}(A) = \nu(A \cap N^c)$, and "suffer" (i.e. check that it all works out). □

Proposition - The Lebesgue decomposition is unique.

Proof - Suppose that $\nu = \nu_S^{(1)} + \nu_{AC}^{(1)} = \nu_S^{(2)} + \nu_{AC}^{(2)}$. Now

$$\nu_S^{(1)} - \nu_S^{(2)} = \nu_{AC}^{(2)} - \nu_{AC}^{(1)}$$

The left-hand side is singular with respect to $\mu$. The right-hand side is absolutely continuous with respect to $\mu$. Since these are equal, they must both be equal to zero, so that $\nu_S^{(1)} = \nu_S^{(2)}$ and $\nu_{AC}^{(1)} = \nu_{AC}^{(2)}$. □

Corollary (Radon-Nikodym-Lebesgue Decomposition) - Let $\mu, \nu$ be positive, $\sigma$-finite measures. Then there is unique $g : X \to [0, \infty)$ and unique $\nu_S$ such that

$$d\nu = d\nu_S + g\, d\mu, \quad \nu_S \perp \mu$$

Application of Radon-Nikodym: The Dual of $L^p$

Let $X$ be a Banach space. We denote by $X^*$ the *(continuous) dual* of $X$, defined by

$$X^* = \{x^* \mid x^* : X \to \mathbb{R} \text{ linear and continuous}\}$$

Theorem - Let $p \in [1, \infty)$, $\mu$ a positive $\sigma$-finite measure, $\frac{1}{p} + \frac{1}{q} = 1$. Then $L^q$ is isometric to $(L^p)^*$.
Remark - Given $g \in L^q$, define $\Lambda_g : L^p \to \mathbb{R}$ by

$$f \overset{\Lambda_g}{\to} \int_X fg\, d\mu$$

By Hölder, this is well-defined. We'll prove the theorem by showing that $g \mapsto \Lambda_g$ maps $L^q$ to $(L^p)^*$ isometrically.

**Lecture 24 - 10/24/2014**

Theorem - Let $p \in [1, \infty)$, $\mu$ a $\sigma$-finite positive measure, $\frac{1}{p} + \frac{1}{q} = 1$. Then $L^q$ is isometrically isomorphic to $(L^p)^*$.
Lemma - Let $U$ be a Banach space, $T : U \to \mathbb{R}$ linear. Then

$$T \text{ is continuous} \iff \exists C \geqslant 0 \text{ s.t. } |Tu| \leqslant C\|u\| \text{ for all } u$$

Definition - We say $T : U \to \mathbb{R}$ is *bounded* if there is $C \geqslant 0$ such that $|Tu| \leqslant C\|u\|$ for all $u \in U$.
Proof (of Lemma) - Say $T$ is bounded. Then $|Tu - Tv| = |T(u - v)| \leqslant C\|u - v\|$. In particular, $T$ is Lipschitz, so $T$ is continuous.

Conversely, suppose $T$ is continuous. Let $\delta > 0$ such that $\|u\| < \delta$ implies $|T(u)| = |T(u) - T(0)| < 1$. Then for all $v \neq 0$, we have

$$|Tv| = \frac{2\|v\|}{\delta}|T(\frac{\delta}{2}\frac{v}{\|v\|})| < \frac{2}{\delta}\|v\|$$

□

Notation - For $u \in U$, denote by $\|u\|_U$ the norm of $u$ in $U$. Similarly, for $u^* \in U^*$, denote by $\|u^*\|_{U*}$ the norm of $u^*$ in $U^*$ (to be defined).
Definition - Given $u^* \in U^*$, define

$$\|u^*\|_{U*} := \sup_{u \in U, \|u\|_U = 1} u^*(u)$$

(Note: This is finite since each $u^*$ is bounded.)
Exercise $\| \cdot \|_{U*}$ is a norm on $U^*$ and $(U^*, \| \cdot \|_{U*})$ is a Banach space.

Proof (of Theorem) - Consider the map $\Lambda : L^q \to (L^p)^*$ which takes $g \in L^q$ to $\Lambda_g \in (L^p)^*$ defined by $\Lambda_g(f) = \int fg$. The map $\Lambda_g$ is clearly linear. Moreover, it is continuous by Hölder's inequality because $|\Lambda_g(f)| \leqslant \|g\|_q\|f\|_p$ (so it is bounded). Thus, for all $g \in L^q$, $\Lambda_g \in (L^p)^*$.

It remains to show that $\Lambda$ is a linear bijective isometry. Linearity is clear. For isometry, we want $\|\Lambda_g\|_{(L^p)^*} = \|g\|_{L^q}$ for all $g \in L^q$. Note,

$$\|\Lambda_g\|_{(L^p)^*} = \sup_{f \in L^p, f \neq 0} \frac{|\Lambda_g(f)|}{\|f\|_p} = \sup_{f \in L^p, f \neq 0} \frac{1}{\|f\|_p}|\int fg| = \|g\|_q$$

(by the duality lemma from long ago). Thus, $\Lambda$ is an isometry.

Since $\Lambda$ is an isometry, it is automatically injective. We need to show $\Lambda$ is surjective. Let $\phi \in (L^p)^*$. Want $g \in L^q$ such that $\phi = \Lambda_g$.

**Case I:** Suppose $\mu$ is finite. The main idea is to note that for any $A \in \Sigma$, $\chi_A \in L^p$, so $\phi(\chi_A)$ is defined. Define $\nu(A) = \phi(\chi_A)$. We show that

1. $\nu$ is a signed, finite measure.
2. $\nu \ll \mu$, so that by Radon-Nikodym, $d\nu = g\, d\mu$
3. $\phi = \Lambda_g$

For (1), note that $\nu(\varnothing) = \phi(0) = 0$. Given $(A_n) \subseteq \Sigma$ countable and pairwise disjoint, we have for each $N$

$$\nu(\bigcup_{1 \leqslant i \leqslant N} A_i) = \phi(\chi_{\bigcup_{1 \leqslant i \leqslant N} A_i}) = \sum_{1 \leqslant i \leqslant N} \phi(\chi_{A_i}) = \sum_{1 \leqslant i \leqslant N} \nu(A_i)$$

Thus, we have finite additivity for $\nu$. For countable additivity, it's enough to show that

$$\sum_{1 \leqslant i \leqslant N} \chi_{A_i} \overset{N \to \infty}{\to} \chi_{\bigcup_{i \geqslant 1} A_i} \text{ in } L^p$$

Write

$$\|\chi_{\bigcup\limits_{i\geqslant 1} A_i} - \sum_{1\leqslant i\leqslant N} \chi_{A_i}\|_p^p = \|\chi_{\bigcup\limits_{i\geqslant N+1} A_i}\|_p^p = \sum_{i\geqslant N+1} \mu(A_i) \to 0$$

Thus, $\nu$ is countably additive.

For (2), suppose $\mu(A) = 0$. Then $\chi_A = 0$ in $L^q$, so $\nu(A) = \phi(\chi_A) = 0$. By Radon-Nikodym, there is $g$ such that $d\nu = g\, d\mu$.

For (3), note that for any simple function $s$, we have

$$\Lambda_g(s) = \int sg\, d\mu = \int s\, d\nu = \sum_i a_i \nu(A_i) = \phi(s),$$

To show that $g \in L^q$, let $f \in L^p$ be arbitrary. Let $(s_n) \subseteq L^p$ with $s_n \xrightarrow{L^p} f$. Then by continuity of $\phi$, $\phi(s_n) \to \phi(f)$ in $\mathbb{R}$. But $\phi(s_n) = \int s_n\, d\nu \to \int f\, d\nu$ (definition of $\int$). Hence for all $f \in L^p$,

$$\phi(f) = \int f\, d\nu = \int fg\, d\mu = \Lambda_g(f)$$

Now $\phi$ bounded implies $g \in L^q$, since

$$\sup_{\|f\|_p = 1} \int fg\, d\mu = \|\phi\|_{(L^p)^*} < \infty$$

**Case II**: If $\mu$ is $\sigma$-finite, reduce to Case I. □

**Lecture 25 - 10/27/2014**

Last Time - If $p \in [1, \infty)$, $\mu(X) < \infty$, and $\frac{1}{p} + \frac{1}{q} = 1$, there is a bijective isometry $(L^p(X))^* \simeq L^q(X)$.
Remark - Why is this useful? Many important normed vector spaces of functions, e.g. $(L^p(X), \|\cdot\|_p)$, are infinite-dimensional. As a result, we don't get compactness of the closed unit ball (contrast with $\mathbb{R}^n$). The Banach-Alaoglu theorem (from functional analysis) says that the closed unit ball of the dual space $X^*$ is weak-$*$ compact (compact with respect to something called the weak-$*$ topology). In particular, we have some sort of compactness for $L^q(X)$, since it is the dual space of $L^p(X)$.
Remark - For $p = \infty$, $(L^\infty)^* \simeq \{\nu \mid \nu$ a finitely additive measure, $\nu \ll \mu\}$, a "boring space". In search of a better result, let's consider a subspace of $L^\infty$ (which will have a bigger, hopefully more interesting space as its dual).

Suppose that $X$ is a compact metric space, so now $C(X) \subseteq L^\infty(X)$. (Note: $(C(X), \|\cdot\|_\infty)$ is a Banach space since a uniform limit of continuous functions is continuous.) Let $\mu$ be a finite, signed, regular Borel measure on $X$. Define

$$\Lambda_\mu \in C(X)^*, \quad \Lambda_\mu(f) = \int_X f \, d\mu \text{ for all } f \in C(X)$$

Theorem (Riesz) - Denote $\mathcal{M} = \{\mu \mid \mu$ a finite, signed, regular Borel measure on $X\}$. Then the map $\Lambda : \mathcal{M} \to C(X)^*$ defined by

$$(\Lambda(\mu))(f) = \Lambda_\mu(f) = \int_X f \, d\mu$$

is a bijective linear isometry.
Proof - Linearity is immediate. To show that $\mu$ is an isometry, first write

$$\|\mu\| = |\mu|(X) = \mu(P) - \mu(N), \qquad \|\Lambda_\mu\| = \sup_{\|f\|_\infty = 1} |\Lambda_\mu(f)|,$$

where $(P, N)$ is a Hahn decomposition of $X$ with respect to $\mu$, $P \cap N = \varnothing$. Clearly $\|\mu\| \geqslant \|\Lambda_\mu\|$. To get the converse, we'd like to choose $f = \chi_P - \chi_N$ above, but this may not be continuous, so use regularity/Lusin's theorem to approximate $f$.

It remains to show that $\mu \mapsto \Lambda_\mu$ is bijective. Since this map is an isometry, it is injective, so only surjectivity remains.

Suppose $I \in (C(X)^*)$ ("$I$ for integral"). We need to find $\mu \in \mathcal{M}$ such that $I(f) = \int_X f \, d\mu$ for all $f \in C(X)$.

**Case I:** Suppose $I$ is positive. That is, given $f \in C(X)$ with $f \geqslant 0$, we have $I(f) \geqslant 0$ (intuitively, $I$ looks like integration with respect to a positive measure). For any open $U \subseteq X$, define

$$\mu^*(U) = \sup\{I(f) \mid f \in C(X),\ 0 \leqslant f \leqslant 1, \operatorname{supp}(f) \subseteq U\}$$

(We're trying to recover the measure of any Borel set, but our only means of approximation is continuous functions.) Now for any $A \subseteq X$, define

$$\mu^*(A) = \inf\{\mu^*(U) \mid U \supseteq A \text{ is open}\}$$

It's not hard to check that $\mu^*$ is indeed an outer measure, so define, as in Carathéodory's theorem, $\Sigma = \{E \mid \mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$ for all $A \subseteq X\}$. Now $\mu^* \restriction_\Sigma$ is a measure.

**Lemma 1** $\Sigma \supseteq \mathcal{B}(X)$.

**Lemma 2** For all $f \in C(X)$, $I(f) = \int_X f \, d\mu$.

**Proof (Lemma 1)** It's enough to show $U \in \Sigma$ for all $U \subseteq X$ open. Let $A \subseteq X$. We need to show that $\mu^*(A) \geqslant \mu^*(A \cap U) + \mu^*(A \cap U^c)$.

- If $A$ is open, write $A = V$. Fix $\epsilon > 0$. Let $f, g \in C(X)$ such that $0 \leqslant f, g \leqslant 1$, and
$$\text{supp}(f) \subseteq V \cap U, \quad I(f) \geqslant \mu^*(V \cap U) - \tfrac{\epsilon}{2}$$
$$\text{supp}(g) \subseteq V \backslash \text{supp}(f), \quad I(g) \geqslant \mu^*(V \backslash \text{supp}(f)) - \tfrac{\epsilon}{2} \geqslant \mu^*(V \cap U^c) - \tfrac{\epsilon}{2}$$
Now $f + g \in C(X)$, $0 \leqslant f + g \leqslant 1$, and $\text{supp}(f + g) \subseteq V$, so
$$\mu^*(V) \geqslant I(f + g) = I(f) + I(g) \geqslant \mu^*(V \cap U) + \mu^*(V \cap U^c) - \epsilon$$
Taking the limit $\epsilon \to 0^+$, we are done.
- Let $A$ be arbitrary. Fix $\epsilon > 0$ and let $A \subseteq V \subseteq X$ open with $\mu^*(V) \leqslant \mu^*(A) + \epsilon$. Now
$$\mu^*(A) + \epsilon \geqslant \mu^*(V) = \mu^*(V \cap U) + \mu^*(V \cap U^c) \geqslant \mu^*(A \cap U) + \mu^*(A \cap U^c)$$
Taking $\epsilon \to 0^+$, we are done.

□

**Proof (Lemma 2)** Need to show $f \in C(X)$ implies $I(f) = \int_X f \, d\mu$.

- Step 1: Suppose $f \geqslant \chi_A$, where $A \in \mathcal{B}(X)$. We show that $I(f) \geqslant \mu(A)$. Given $\epsilon > 0$, define $U = \{f > 1 - \epsilon\}$. Note that $U$ is open and $U \supseteq A$. If $g \in C(X)$ with $0 \leqslant g \leqslant 1$ and $\text{supp}(g) \subseteq U$, then
$$g \leqslant \tfrac{f}{1-\epsilon} \implies I(g) \leqslant \tfrac{1}{1-\epsilon} I(f),$$
since we are assuming $I(\cdot)$ is a positive operator. Taking the supremum over all such $g$, we have $\mu(U) \leqslant \frac{I(f)}{1-\epsilon}$. Taking the limit $\epsilon \to 0^+$, we have $\mu(A) \leqslant I(f)$. (Note: Although $U = U(\epsilon)$, we clearly have $A \subseteq U(\epsilon)$ for any $\epsilon > 0$, so $\mu(A) \leqslant \mu(U(\epsilon))$.)

*To be continued...*

**Lecture 26 - 10/29/2014**

We continue with the proof of Riesz's theorem (refer to previous lecture's notes).

**Lemma 2** $I(f) = \int_X f\, d\mu$ for all $f \in C(X)$.

**Proof**
- Step 1: We showed that $\chi_A \leqslant f$ implies $\mu(A) \leqslant I(f)$ for $A \in \mathcal{B}(X)$.
- Step 2: Given $s = \sum_i a_i \chi_{A_i} \geqslant 0$, where $\{A_i\}$ are disjoint, we show that

$$s \leqslant f \implies \int_s d\mu \leqslant I(f)$$

To do this, fix $\epsilon > 0$. For each $i$ select $K_i \subseteq A_i$ compact, $U_i \supseteq K_i$ open such that the sets $\{U_i\}$ are disjoint and $\frac{f}{1-\epsilon} \geqslant a_i$ on $U_i$. (Note: Disjointness is possible since disjoint compact subsets of a metric space can be separated.) Now let $\phi_i : X \to [0,1]$ be continuous such that $\phi_i = 1$ on $K_i$ and $\operatorname{supp}(\phi_i) \subseteq U_i$. For each $i$, we have

$$a_i \chi_{K_i} \leqslant \tfrac{\phi_i f}{1-\epsilon} \implies a_i \mu(K_i) \leqslant \tfrac{I(\phi_i f)}{1-\epsilon}$$

Summing over $i$, we have

$$\sum_i a_i \mu(K_i) \leqslant \sum_i \tfrac{I(\phi_i f)}{1-\epsilon} = \tfrac{I((\sum_i \phi_i) f)}{1-\epsilon} \leqslant \tfrac{I(f)}{1-\epsilon},$$

since the supports of the functions $\{\phi_i\}$ are disjoint. Taking the supremum over compact $K_i \subseteq A_i$, we have

$$\int s\, d\mu \leqslant \tfrac{I(f)}{1-\epsilon}$$

Taking the limit $\epsilon \to 0^+$, we have $\int s\, d\mu \leqslant I(f)$, as desired.
- Step 3: Exercise: If $0 \leqslant f \leqslant t$ and $t$ is simple, then $I(f) \leqslant \int_X t\, d\mu$. (Appeal to Step 2 and finiteness of $\mu$.)
- Step 4: By the Exercise, given $f \in C(X)$, we have

$$I(f^+) \geqslant \sup_{\substack{0 \leqslant s \leqslant f^+ \\ s \text{ simple}}} \int_X s\, d\mu = \int_X f^+\, d\mu = \inf_{\substack{0 \leqslant f^+ \leqslant t \\ t \text{ simple}}} \int_X t\, d\mu \geqslant I(f^+)$$

which implies $\int_X f^+\, d\mu = I(f^+)$, so $\int_X f\, d\mu = I(f)$.

**Case II:** Let $I \in C(X)^*$ be arbitrary. Given $f \geqslant 0$, define $I^+(f) = \sup\{I(g) \mid 0 \leqslant g \leqslant f,\ g \in C(X)\}$. Note $I^+(f) \geqslant 0$ (take $g = 0$). Moreover, this is finite since $I$ is bounded. Check that this is linear. Define

$$I^-(f) = \inf\{I(g) \mid 0 \leqslant g \leqslant f,\ g \in C(X)\}$$

or alternately define $I^- = I^+ - I$ (we want $I = I^+ - I^-$). Now for $f \in C(X)$ arbitrary, define

$$I^+(f) = I^+(f^+) - I^+(f^-), \qquad I^-(f) = I^+(f) - I(f)$$

Check that everything works out! We get $I^+ \in (C(X))^*$ by applying Case I, same for $I^-$.

This finishes the proof. □

Remark - The decomposition $I = I^+ - I^-$ corresponds to the Jordan decomposition. Explicitly, if $I = \mu$, then $I^+ = \mu^+$ and $I^- = \mu^-$.

Two Extensions of the Theorem

1. If $X$ is *locally compact* and $I : C_c(X) \to \mathbb{R}$ is positive, then there is a unique regular Borel measure $\mu$ such that $I(f) = \int_X f \, d\mu$ for all $f \in C_c(X)$.

2. If $X$ is locally compact, let $\mathcal{M}$ be the set of finite, signed, regular Borel measures on $X$. We have $C_c(X) \subseteq L^\infty(X)$, so we can define $C_0(X) = \overline{C_c(X)}$ in $L^\infty(X)$. We call $C_0(X)$ the *set of continuous functions that vanish at infinity.* It can be shown that $(C_0(X))^* = \mathcal{M}$, where the map

$$\Lambda : \mathcal{M} \to (C_0(X))^*, \qquad \Lambda_\mu(f) = \int_X f \, d\mu$$

is a bijective linear isometry.

**Lecture 27 - 10/31/2014**

Exercise: Continuity of Regular Measures - Let $X$ be a compact metric space, $\mu$ a finite regular positive Borel measure on $X$ such that $\mu(\{x\}) = 0$ for all $x \in X$. Then for any $\alpha \in (0, \mu(X))$, there is $A \subseteq X$ with $\mu(A) = \alpha$. (Hint: Use Riesz.)

**Product Measures and Fubini's Theorem**

Setting - $(X, \Sigma, \mu)$ and $(Y, \tau, \nu)$ are measure spaces, $\mu$, $\nu$ $\sigma$-finite.
Definition - We denote $\Sigma \times \tau = \{A \times B \mid A \in \Sigma, B \in \tau\}$, the set of all "rectangles" in $A \times B$. The *product* $\sigma$-algebra on $A \times B$ is defined by

$$\Sigma \otimes \tau := \sigma(\Sigma \times \tau)$$

Theorem - There is a unique measure $\pi$ on $(X \times Y, \Sigma \otimes \tau)$ such that

$$\pi(A \times B) = \mu(A)\nu(B)$$

for all $A \in \Sigma$, $B \in \tau$. We call $\pi$ the *product measure* on $(X \times Y, \Sigma \otimes \tau)$.
Remark - This fails if $\mu$ or $\nu$ is not $\sigma$-finite.

We begin by stating a series of theorems which we will then proceed to prove. We take this approach because the final result (Fubini's theorem) plays a decisive role in our method of proof for the first result (existence of the product measure).

Theorem - Let $f : X \times Y \to [0, \infty]$ be $\Sigma \otimes \tau$-measurable.

(a) For all $x \in X$, the map $y \mapsto f(x, y)$ is $\tau$-measurable. Similarly, for all $y \in Y$, the map $x \mapsto f(x, y)$ is $\Sigma$-measurable.

(b) The functions $x \mapsto \int_Y f(x, y)\, d\nu(y)$ and $y \mapsto \int_X f(x, y)\, d\mu(x)$ are $\Sigma$- and $\tau$-measurable respectively.

(c) (Tonelli)

$$\int_{X \times Y} f\, d\pi = \int_X \Big( \int_Y f(x, y)\, d\nu(y) \Big) d\mu(x) = \int_Y \Big( \int_X f(x, y)\, d\mu(x) \Big) d\nu(y) \qquad (*)$$

Theorem (Fubini) - Suppose $f \in L^1(X \times Y, \pi)$.

(a) For a.e. $x \in X$, $\int_Y f(x, y)\, d\nu(y)$ is finite and $\mu$-integrable.

(b) For a.e. $y \in Y$, $\int_X f(x, y)\, d\mu(x)$ is finite and $\nu$-integrable.

(c) The equality $(*)$ holds and the integrals are finite.

Proof (existence of product measure)
Main Idea - Given $E \subseteq X \times Y$, for each $x \in X$ define $S_x(E) = \{y \in Y \mid (x, y) \in E\}$. For each $y \in Y$, define $T_y(E) = \{x \in X \mid (x, y) \in E\}$. Now *define*

$$\pi(E) = \int_X \nu(S_x(E))\, d\mu(x)$$

We need to show three things:

1. $S_x(E) \in \tau$ for all $x \in X$.

2. The function $x \mapsto \nu(S_x(E))$ is $\Sigma$-measurable.

3. $\pi$ is a measure.

First, we prove uniqueness. Say $\pi_1$, $\pi_2$ are two such measures. If $\mu, \nu$ are finite, then the set $\{\pi_1 = \pi_2\}$ is a $\lambda$-system containing $\Sigma \times \tau$, which is a $\pi$-system. By Dynkin's theorem, $\{\pi_1 = \pi_2\} = \Sigma \otimes \tau$. If $\mu, \nu$ are $\sigma$-finite, then take limits and use the previous case.

For existence, we prove a sequence of lemmas.

**Lemma 1** For all $E \in \Sigma \otimes \tau$, $x \in X$, $y \in Y$, $S_x(E) \in \tau$ and $T_y(E) \in \Sigma$.

**Proof** Define $\mathcal{F} = \{E \in \Sigma \otimes \tau \mid \forall x \in X,\ S_x(E) \in \tau\}$. This is a $\sigma$-algebra containing the rectangles $\Sigma \times \tau$, so $\mathcal{F} = \Sigma \otimes \tau$. The argument is the same for $T_y(E)$. □

**Lemma 2** For all $E \in \Sigma \otimes \tau$, the function $x \mapsto \nu(S_x(E))$ is $\Sigma$-measurable. Similarly, the function $y \mapsto \mu(T_y(E))$ is $\tau$-measurable.

**Proof** Denote $\Lambda = \{E \in \Sigma \otimes \tau \mid x \mapsto \nu(S_x(E)) \text{ is } \Sigma\text{-measurable}\}$. This is not obviously a $\sigma$-algebra (because of the problem of non-disjoint unions). Instead, we show that $\Lambda$ is a $\lambda$-system! Since $\Lambda$ contains the rectangles $\Sigma \times \tau$, we then have $\Lambda = \Sigma \otimes \tau$.

First suppose that $\mu, \nu$ are finite.

(1) Clearly $X \times Y \in \Lambda$, since $\nu(S_x(X \times Y)) = \nu(Y)$, a constant function.

(2) Suppose $A, B \in \Lambda$, $A \subseteq B$. We have

$$\nu(S_x(B \backslash A)) = \nu(S_x(B)) - \nu(S_x(A)),$$

a difference of measurable functions, which is measurable. Thus $B \backslash A \in \Lambda$.

(3) Let $(A_n) \subseteq \Lambda$ with $A_n \subseteq A_{n+1}$. We have

$$\nu(S_x(\bigcup_n A_n)) = \nu(\bigcup_n S_x(A_n)) = \lim_{n \to \infty} \nu(S_x(A_n)),$$

a limit of measurable functions, which is measurable. Thus $\bigcup_n A_n \in \Lambda$. Hence, $\Lambda$ is a $\lambda$-system.

If $\mu, \nu$ are $\sigma$-finite, take limits and use the previous case. □

*To be continued...*

**Lecture 28 - 11/03/2014**

Quotes (A bowl of candy is sitting at the front of the classroom) Student: "Are we going over the Reese's Representation Theorem with candy?" Professor: "What's the connection between Riesz and candy?" *(slaps forehead)* "I'm slow."
Proof (of existence of product measure, continued) - We still need to show that $\pi(E) = \int_X \nu(S_x(E))\, d\mu(x)$ is a measure. Let $(E_n) \subseteq \Sigma \otimes \tau$ be pairwise-disjoint. Now

$$\begin{aligned}\pi(\bigcup_n E_n) &= \int_X \nu(S_x(\bigcup_n E_n))\, d\mu(x)\\ &= \int_X \sum_n \nu(S_x(E_n))\, d\mu(x)\\ &= \sum_n \int_X \nu(S_x(E_n))\, d\mu(x) \quad \text{(Beppo-Levi/MCT)}\\ &= \sum_n \pi(E_n)\end{aligned}$$

Hence, $\pi$ is a measure. Moreover, for any $A \times B \in \Sigma \times \tau$, we have

$$\pi(A \times B) = \int_X \nu(S_x(A \times B))\, d\mu(x) = \int_X \nu(B)\chi_A(x)\, d\mu(x) = \nu(B)\int_A d\mu(x) = \mu(A)\nu(B)$$

□

Remark - We could have defined the product measure more constructively, starting with some outer measure that uses coverings of $E \in \Sigma \otimes \tau$ by countable collections of rectangles $(A_n \times B_n) \subseteq \Sigma \times \tau$. Then do some sort of Carathéodory argument. There are two drawbacks to this: First of all, this would be very tedious. Second of all, this construction makes it difficult to recover the "nicer" definition of $\pi$ in terms of integration. We will see that with our definition of $\pi$, it doesn't take too much work to prove Tonelli's and Fubini's theorems.

Proof (Tonelli) - Let $f : X \times Y \to [0, \infty]$ be $\Sigma \otimes \tau$-measurable. We need to show that

(a) For all $x \in X$, $y \in Y$, the maps $f(x, \cdot)$ and $f(\cdot, y)$ are $\tau$-measurable, $\Sigma$-measurable respectively.

(b) The maps $\int_Y f(\cdot, y)\, d\nu(y)$ and $\int_X f(x, \cdot)\, d\mu(x)$ are $\Sigma$- and $\tau$-measurable respectively.

(c) (Tonelli)

$$\int_{X\times Y} f\, d\pi = \int_X \Big(\int_Y f(x, y)\, d\nu(y)\Big)d\mu(x) = \int_Y \Big(\int_X f(x, y)\, d\mu(x)\Big)d\nu(y) \qquad (*)$$

We proceed in several steps.

1. Say $f = \chi_E$, where $E \in \Sigma \otimes \tau$. By definition, we have

$$\int_{X\times Y} f\, d\pi = \pi(E) = \int_X \nu(S_x(E))\, d\mu(x) = \int_X \Big(\int_Y \chi_E(x, y)\, d\nu(y)\Big)d\mu(x)$$

   and we have already shown that (a) and (b) hold. By symmetry, the function $\tilde{\pi}(E) = \int_Y \mu(T_y(E))\, d\nu(y)$ is a measure that satisfies $\tilde{\pi}(A \times B) = \mu(A)\nu(B)$. By uniqueness of the product measure, we have $\pi = \tilde{\pi}$, so $(*)$ holds.

2. By linearity and the previous step, $(*)$ holds if $f$ is a positive simple function. Moreover, (a) and (b) are clearly preserved under positive linear combinations.

3. Now let $f : X \times Y \to [0, \infty]$ be arbitrary $\Sigma \otimes \tau$-measurable. Let $s_n \nearrow f$ be a sequence of positive, simple, $\Sigma \otimes \pi$-measurable functions. By MCT, (a) and (b) clearly hold. For $(*)$, write

$$\begin{aligned}\int_{X\times Y} f\, d\pi \overset{\text{(MCT)}}{=} \lim_{n\to\infty} \int_{X\times Y} s_n\, d\pi &= \lim_{n\to\infty} \int_X \int_Y s_n(x,y)\, d\nu(y)\, d\mu(x) \\ &= \int_X \Big( \lim_{n\to\infty} \int_Y s_n(x,y)\, d\nu(y)\Big) d\mu(x) \quad \text{(MCT)} \\ &= \int_X \int_Y \lim_{n\to\infty} s_n(x,y)\, d\nu(y)\, d\mu(x) \quad \text{(MCT)} \\ &= \int_X \int_Y f\, d\nu(y)\, d\mu(x)\end{aligned}$$

□

Proof (Fubini) - Let $f \in L^1(X \times Y, \pi)$. Write $f = f^+ - f^-$. Now

$$\begin{aligned}\int_{X\times Y} f\, d\pi &= \int_{X\times Y} (f^+ - f^-)\, d\pi \\ &= \int_{X\times Y} f^+\, d\pi - \int_{X\times Y} f^-\, d\pi \\ &= \int_X \int_Y f^+(x,y)\, d\nu(y)\, d\mu(x) - \int_X \int_Y f^-(x,y)\, d\nu(y)\, d\mu(x) \quad \text{(Tonelli)} \\ &= \int_X \Big( \int_Y f^+\, d\nu(y) - \int_Y f^-\, d\nu(y)\Big) d\mu(x) \qquad (**) \\ &= \int_X \int_Y f(x,y)\, d\nu(y)\, d\mu(x)\end{aligned}$$

Why does step $(**)$ hold? This is because

$$\mu(\{x \mid \int_Y f^+(x,y)\, d\nu(y) = \infty\}) = 0, \qquad \mu(\{x \mid \int_Y f^-(x,y)\, d\nu(y) = \infty\}) = 0$$

In particular, $\int_Y f^+(x,y)\, d\nu(y) - \int_Y f^-(x,y)\, d\nu(y)$ is well-defined except on $\mu$-null set. □

Application of Fubini/Tonelli - Let $f : X \to [0, \infty)$ be measurable, $\mu$ $\sigma$-finite on $X$. Then

1. $\int_X f\, d\mu = \int_0^\infty \mu(\{f > x\})\, d\lambda(x)$
2. If $\phi : [0, \infty) \to [0, \infty)$ is $C^1$, increasing, and satisfies $\phi(0) = 0$, then

$$\int_X \phi(f)\, d\mu = \int_0^\infty \mu(\{f > x\})\phi'(x)\, d\lambda(x)$$

Note that (1) implies (2), since (1) implies

$$\int_X \phi(f) = \int_0^\infty \mu(\{\phi(f) > x\})\, d\lambda(x) = \int_0^\infty \mu(\{f > x\})\phi'(x)\, d\lambda(x)$$

by change of variables. To prove (1), define $E = \{(x, \lambda) \mid \lambda \in [0, f(x)]\}$. Then by Tonelli,

$$\int_{X\times\mathbb{R}} \chi_E\, d\pi = \int_X \int_0^\infty \chi_E\, d\lambda\, d\mu(x) = \int_X f(x)\, d\mu(x)$$

$$\int_{X\times\mathbb{R}} \chi_E\, d\pi = \int_0^\infty \int_X \chi_E\, d\lambda\, d\mu(x) = \int_0^\infty \mu(\{f \geqslant y\})\, d\lambda(y)$$

Remark - We also have $\int f = \int_0^\infty \mu(\{f > y\})\, d\lambda(x)$, since we're only cutting out the graph of the function (which has measure zero).

**Lecture 29 - 11/05/2014**

Setting - All functions are functions on $\mathbb{R}^d$.
Motivation - Consider a continuous but erratic function $f(t)$, e.g. the price of a stock over the course of a year, which typically moves up and down quite a lot. Instead of studying $f(t)$, we can study a "moving average" of $f(t)$, defined by $F(t) = \int_{t-\epsilon}^{t+\epsilon} f(t) \cdot \frac{1}{2\epsilon}\, dt$ (note: $F(t)$ is the average value of the function $f(t)$ over the interval $(t-\epsilon, t+\epsilon)$). This function generally looks smoother and is easier to interpret.

Definition - Let $f, g$ be functions on $\mathbb{R}^d$. We define a new function $f * g$ ("$f$ star $g$" or "$f$ convolved with $g$") by

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x-y)g(y)\, dy$$

Remark - By making a change of variables, we see that $f * g = g * f$.
Remark - In the Motivation above, $F(t) = f * (\frac{1}{2\epsilon}\chi_{[-\epsilon,\epsilon]})$.

When is $f * g$ defined? To answer this, we must make assumptions on $f$ and $g$. For simplicity, let's suppose first that $f, g \in L^1$. Then by Tonelli's theorem, we have

$$\begin{aligned}\int_{\mathbb{R}^d} |f| * |g|\, dx &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(x-y)||g(y)|\, dydx \\ &= \int_{\mathbb{R}^d} |g(y)| \int_{\mathbb{R}^d} |f(x-y)|\, dxdy \\ &= \int_{\mathbb{R}^d} |g(y)|\|f\|_1\, dy \\ &= \|f\|_1\|g\|_1\end{aligned}$$

In particular, $\int_{\mathbb{R}^d} |f| * |g|\, dx$ is finite, so not only is $f * g$ defined a.e., but $f * g \in L^1$.

The previous result is nice. How does it generalize? That is, given $f \in L^p$ and $g \in L^q$, is $f * g$ defined? Can we find $r$ such that $f * g \in L^r$?

We seek an inequality of the form $\|f * g\|_r \leqslant \|f\|_p\|g\|_q$. As we've done previously, let's try "counting dimensions". View $f, g$ as dimensionless functions, $\ell$ denotes length, so that $\mathbb{R}^d$ has dimension $\ell^d$. We note that:

- $f * g$ has dimension $\ell^d$, so then $\|f * g\|_r$ has dimension $\ell^{d+d/r}$.
- $\|f\|_p\|g\|_q$ has dimension $\ell^{d/p}\ell^{d/q}$.

If $f * g \in L^r$, then we should expect $d + \frac{d}{r} = \frac{d}{p} + \frac{d}{q}$, or $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. This agrees with the result we just derived above; take $p = q = r = 1$.

Theorem (Young's Inequality) - Let $p, q \in [1, \infty]$. If $f \in L^p$, $g \in L^q$, then $f * g \in L^r$, where $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. In fact,

$$\|f * g\|_r \leqslant \|f\|_p\|g\|_q$$

Proof - Let $r'$ be the conjugate of $r$. By the duality lemma, it's enough to show that for all $h \in L^{r'}$, we have

$$\int (f * g)h \leqslant \|f\|_p\|g\|_q\|h\|_{r'} \quad (*)$$

For simplicity, assume $f, g, h \geqslant 0$. Otherwise, replace $f, g, h$ by $|f|, |g|, |h|$ and use monotonicity of the integral. Let $p', q'$ be the conjugates of $p, q$ respectively. We have

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}, \quad \frac{1}{p} = 1 - \frac{1}{p'}, \quad \frac{1}{q} = 1 - \frac{1}{q'} \implies \frac{1}{p'} + \frac{1}{q'} + \frac{1}{r} = 1 \qquad (**)$$

In light of (∗∗), we would like to apply Hölder's inequality (the general version for $n$ functions, $n \geqslant 2$), followed by Tonelli's theorem. To start, write

$$\begin{aligned}\int_{\mathbb{R}^d}(f*g)h &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(x-y)g(y)h(x)\;dydx \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(f(x-y)^{\alpha}g(y)^{\beta})\cdot(f(x-y)^{1-\alpha}h(y)^{\gamma})\cdot(g(y)^{1-\beta}h(x)^{1-\gamma})\;dydx\end{aligned}$$

*(Note: Applying Hölder's inequality will separate the three functions $f, g, h$ into three pairs of two functions, which will then permit us to use Tonelli's theorem.)* We'd like one of each of the three factors above to be in $L^{p'}$, $L^{q'}$, and $L^r$. We know $f \in L^p$, $g \in L^q$, $h \in L^{r'}$, so it's enough to find $\alpha, \beta, \gamma$ such that

$$\alpha r = p,\ \beta r = q, \qquad (1-\alpha)q' = p,\ \gamma q' = r', \qquad (1-\beta)p' = q,\ (1-\gamma)p' = r'$$

From the first, second, and fourth equation, we require that $\alpha = \frac{p}{r}$, $\beta = \frac{q}{r}$, $\gamma = \frac{r'}{q'}$. It is straightforward to verify that these values satisfy the third, fifth, and sixth equations. By Hölder, we have

$$\begin{aligned}\int_{\mathbb{R}^d}(f*g)h &\leqslant (\iint f(x-y)^p g(y)^q)^{1/r}(\iint f(x-y)^p h(x)^{r'})^{1/q'}(\iint g(y)^q h(x)^{r'})^{1/p'} \\ &= \|f\|_p^{p/r}\|g\|_q^{q/r}\|f\|_p^{p/q'}\|g\|_{r'}^{r'/q'}\|g\|_q^{q/p'}\|h\|_{r'}^{r'/p'} \\ &= \|f\|_p^{\alpha+(1-\alpha)}\|g\|_q^{\beta+(1-\beta)} + \|h\|_{r'}^{\gamma+(1-\gamma)} \\ &= \|f\|_p\|g\|_q\|h\|_{r'}\end{aligned}$$

This finishes the proof. □

Mollification

Definition - Let $\{\phi_n\}$ be a sequence of functions on $\mathbb{R}^d$. We say $\{\phi_n\}$ is an *approximate identity* if

1. $\phi_n \geqslant 0$ for all $n$.
2. $\int_{\mathbb{R}^d}\phi_n = 1$ for all $n$¿
3. For each $\epsilon > 0$, we have $\lim\limits_{n\to\infty}\int_{|x|>\epsilon}\phi_n(x)\;dx = 0$.

Example - Let $\phi \in L^1$ with $\phi \geqslant 0$, $\int_{\mathbb{R}^d}\phi = 1$. For any $\epsilon > 0$, define $\phi_\epsilon(x) = \frac{1}{\epsilon^d}\phi(\frac{x}{\epsilon})$. Then $\{\phi_\epsilon\}_{\epsilon\to 0}$ is an approximate identity.

Approximate identities are useful because they provide a means to approximate arbitrary functions by a sequence of much nicer functions. For example, given an approximate identity consisting of smooth functions, we can use these to approximate many non-smooth functions by smooth functions. In particular...

Proposition - Let $\{\phi_n\}$ be an approimate identity.

1. If $f \in L^p$ for some $p \in [1, \infty)$, then $f * \phi_n \xrightarrow{L^p} f$, i.e. $\|f*\phi_n - f\|_p \to 0$.
2. If $f \in L^\infty$ and $f$ is continuous at some $x \in \mathbb{R}^d$, then $f * \phi_n(x) \to f(x)$.

Proof - We prove (2), leaving (1) for next time. Fix $\epsilon > 0$. Write

$$\begin{aligned} |\phi_n * f(x) - f(x)| &= |\int \phi_n(y) f(x-y)\, dy - f(x)| \\ &= |\int \phi_n(y)(f(x-y) - f(x))\, dy| \\ &\leqslant \int \phi_n(y)|f(x-y) - f(x)|\, dy \\ &= \int_{|y|<\delta} (\cdots)\, dy + \int_{|y|\geqslant\delta} (\cdots)\, dy \end{aligned}$$

Pick $\delta > 0$ sufficiently small so that $|f(x-y) - f(x)| < \frac{\epsilon}{2}$ for all $|y| < \delta$. This bounds the first term by $\frac{\epsilon}{2}$. Now let $N$ be sufficiently large so that $n \geqslant N$ implies $\int_{|y|\geqslant\delta} \phi_n(y)\, dy \leqslant \frac{\epsilon}{4\|f\|_\infty}$. This finishes the proof of (2). $\square$

**Lecture 30 - 11/07/2014**

Remark - Given a measure $\mu$ on $\mathbb{R}^d$, define $f * \mu(x) = \int_{\mathbb{R}^d} f(x-y)\, d\mu(y)$. For example, we have $f * \delta_0(x) = f(x)$ (i.e. $\delta_0$ is the *identity* for this operation $*$). This justifies the nomenclature "approximate identity", since $f * \phi_n \to f = f * \delta_0$.

Recall - From last time, we still need to prove that $f \in L^p$, $p \in [1, \infty)$ implies $f * \phi_n \xrightarrow{L^p} f$.
Proof - We make use of two facts from previous homeworks. First, that if $p \in [1, \infty)$ and $f \in L^p$, then $\tau_y f \xrightarrow{L^p} f$ as $|y| \to 0$ (here, $\tau_y f(x) = f(x-y)$). The second is Minkowski's inequality for integrals. Fix $\epsilon > 0$ and write

$$f * \phi_n(x) - f(x) = \int f(x-y)\phi_n(y)\, dy - f(x) = \int (\tau_y f(x) - f(x))\phi_n(y)\, dy$$

Now

$$\begin{aligned}\|f * \phi_n - f\|_p &= (\int |\int (\tau_y f(x) - f(x))\phi_n(y)\, dy|^p\, dx)^{1/p} \\ &\leqslant \int \|\tau_y f - f\|_p \phi_n(y)\, dy \\ &= \int_{|y|<\delta} \|\tau_y f - f\|_p \phi_n(y)\, dy + \int_{|y|\geqslant\delta} \|\tau_y f - f\|_p \phi_n(y)\, dy\end{aligned}$$

Take $\delta > 0$ small enough so that $\|\tau_y - f\|_p < \frac{\epsilon}{2}$ for all $|y| < \delta$. Then take $N$ sufficiently large so that $\int_{|y|\geqslant\delta} \phi_n(y)\, dy < \frac{\epsilon}{4\|f\|_p}$. This finishes the proof. □

Application: Fourier Series
Quote - "Whatever I say for the first lecture and a half probably is OK with just the Riemann integral, but after that you'll see how much more you can get."
Setup - We consider periodic functions on $[0, 1]$. Denote $e_n(x) = e^{2\pi inx}$, $n \in \mathbb{Z}$. Define

$$L^2_{\text{per}}[0,1] = \{f : \mathbb{R} \to \mathbb{C} \mid f(x+1) = f(x), \quad \int_0^1 |f|^2 < \infty\}$$

If $f, g \in L^2_{\text{per}}$, define $\langle f, g\rangle = \int_0^1 f\overline{g}$. (Note: To integrate a complex-valued function, simply integrate the real and imaginary parts separately, then add them back together. "This is the only sensible way to integrate such functions.") This defines an inner product on $L^2_{\text{per}}$ that induces the norm $\|f\|_2 = \int_0^1 |f|^2$.

Definition - Given $f \in L^2_{\text{per}}[0,1]$, define $\hat{f}(n) = \langle f, e_n\rangle = \int_0^1 f(x)e^{-2\pi inx}\, dx$ (Fourier coefficients).
Definition - Given $N \geqslant 0$, define $S_N f = \sum\limits_{|n|\leqslant N} \hat{f}(n)e_n$ (partial sums of Fourier series).
Remark - The set $\{e_n\}_{n\in\mathbb{Z}}$ is orthonormal, i.e. $\langle e_n, e_m\rangle = \delta_{nm}$.
Goal 1 - Show that $S_N f \xrightarrow{L^2} f$ for all $f \in L^2_{\text{per}}$.

One useful observation is that $\langle S_N f, f - S_N f\rangle = 0$ (easy to check). More generally, we have the following:

Lemma - Let $p_N \in \text{span}\{e_{-N}, \ldots, e_N\}$. Then $\langle f - S_N f, p_N\rangle = 0$.
Proof - By linearity, it's enough to check this for $p_N = e_m$, where $|m| \leqslant N$. Write

$$\langle f, e_m\rangle = \hat{f}(m), \qquad \langle S_N f, e_m\rangle = \sum_{|n|\leqslant N} \hat{f}(n)\langle e_n, e_m\rangle = \hat{f}(m)$$

□

Corollary 1 (Bessel's inqeuality) - $\|\hat{f}\|_{\ell^2} \leqslant \|f\|_{L^2}$
Proof - Set $E_N = f - S_N f$. Then by the Pythagoras theorem for inner products,

$$S_N f + E_N = f, ; \langle S_N f, E_N \rangle = 0 \implies \|f\|_2^2 = \|S_N f\|_2^2 + \|E_N\|_2^2 \geqslant \|S_N f\|_2^2 = \sum_{|n| \leqslant N} |\hat{f}(n)|^2$$

Taking the limit $N \to \infty$, we have the result. □

Corollary 2 - Suppose there is a sequence $(p_N)$ such that for each $N$, $p_N \in \operatorname{span}\{e_{-N}, \ldots, e_N\}$ and $p_N \xrightarrow{L^2} f$. Then $S_N f \xrightarrow{L^2} f$.
Proof - It's enough to show that $\|f - S_N f\|_2 \leqslant \|f - p_N\|$. This follows by basic facts about projections for inner product spaces. Explicitly, write

$$f - p_N = (f - S_N f) + (S_N f - p_N) \implies \|f - p_N\|_2^2 = \|f - S_N f\|_2^2 + \|S_N f - p_N\|_2^2 \geqslant \|f - S_N f\|_2^2,$$

where we have used the fact that $f - S_N f$ and $S_N f - p_N$ are orthogonal (by the lemma). □

Let's relate $S_N f$ back to convolutions and approximate identities. Write

$$S_N f(x) = \sum_{|n| \leqslant N} \hat{f}(n) e^{2\pi i n x} = \sum_{|n| \leqslant N} \int_0^1 f(y) e^{-2\pi i n y}\, dy\, e^{2\pi i n x} = \int_0^1 \sum_{|n| \leqslant N} e^{2\pi i n (x-y)} f(y)\, dy$$

This motivates the definition $D_N(x) = \sum_{|n| \leqslant N} e^{2\pi i n x}$, so that $S_N f = D_N * f$. We call $D_N$ the ($N$th order) *Dirichlet kernel*.

By interpreting $D_N$ as a geometric series, we may express it as a ratio of sines. What we'd like is to say that $\{D_N\}$ is an approximate identity. Unfortunately, the functions $\{D_N\}$ are not even positive! Next try: Maybe $\{|D_N|\}$? This also fails, since $\int |D_N| \approx \ln N$. What can we do?

Instead of considering the partial sums, let's try the Cesàro sums! That is, instead of considering $S_N f$, let's consider

$$\sigma_N f = \tfrac{1}{N} \sum_{n=0}^{N} S_n f,$$

the average of the first $N$ partial sums. We may write

$$\sigma_N f = \tfrac{1}{N} \sum_{n=0}^{N} D_n * f = (\tfrac{1}{N} \sum_{n=0}^{N} D_n) * f$$

Denote $K_N = \frac{1}{N} \sum_{n=0}^{N} D_n$, the *Fejér kernel*.

Fact - $\{K_N\}$ is an approximate identity!
Corollary - $K_N * f \xrightarrow{L^2} f$ for all $f \in L^2_{\text{per}}$.
Corollary - $S_N f \xrightarrow{L^2} f$ for all $f \in L^2_{\text{per}}$.
Quote - "Even though the functions $K_N * f$ are a strictly worse approximation to $f$ than the functions $S_N f$, it's much easier to show that they converge to $f$."

**Lecture 31 - 11/10/2014**

Motivation/Intuition - The functions $\{e_n\}$ are an orthonormal "basis" of $L^2_{\text{per}}$. We expect/hope that $f = \sum_n \hat{f}(n) e_n$.

Remark 1 - Fourier coefficients can be defined much more generally. Certainly the definition extends to $f \in L^1$. In fact, given a finite measure $\mu$ on $[0,1]$, define $\hat{\mu}(n) = \int_0^1 e_n \, d\mu$.

Consequently, if $f \in L^p[0,1]$ for any $p \in [1,\infty]$, we can define $\hat{f}(n) = \int_0^1 f(x)\overline{e_n(x)} \, dx$. With that in mind, we ask:

Question - For which $p \in [1,\infty]$ do we have $S_N f \xrightarrow{L^p} f$?

Answer - For $p = \infty$, this fails. E.g. let $f \in L^\infty$ be discontinuous on a set of positive measure. In fact, it is even possible to find *continuous* $f \in L^\infty$ such that $S_n f \nrightarrow f$ in $L^\infty$. However, on your homework you will show that if $f \in C^\alpha[0,1]$, where $\alpha > 0$, then $S_N f \to f$ in $L^\infty$.

Proposition If $p \in (1,\infty)$, $f \in L^p$, then $S_N f \xrightarrow{L^p} f$.

Proof - Hard harmonic analysis. □

Fact - This result is false for $p = 1$.

Question 2 - We know $f \in L^2_{\text{per}}$ implies $S_N f \xrightarrow{L^2} f$. When do we have pointwise convergence?

Answer - Theorem (Carleson-Hunt) - For $p > 1$, $f \in L^p$, we have $S_N f \to f$ a.e.

* * *

Theorem - The map $L^2_{\text{per}}[0,1] \to \ell^2$, $f \mapsto \hat{f} = (\hat{f}(n))$ is a bijective linear isometry.

Proof - We showed last time that $f \in L^2$ implies $\hat{f} \in \ell^2$, so the map is well-defined. Moreover, since $S_N f \xrightarrow{L^2} f$, then

$$\|\hat{f}\|_{\ell^2} = \lim_{N\to\infty} \sum_{|n|\leqslant N} |\hat{f}(n)|^2 = \lim_{N\to\infty} \|S_N f\|_{L^2} = \|f\|_{L^2}$$

Thus, the map is an isometry, so it is also injective. Linearity is also clear. It remains to prove surjectivity.

Let $(a_n) \in \ell^2$. We'd *like* to define $f = \sum_n a_n e_n$, but is this in $L^2_{\text{per}}$? Yes, because

$$\| \sum_{M\leqslant|n|\leqslant N} a_n e_n \|_{L^2} = ( \sum_{M\leqslant|n|\leqslant N} |a_n|^2 )^{1/2},$$

so the sequence of partial sums is Cauchy, thus $f \in L^2_{\text{per}}$. □

Next Goal - We seek to establish the following correspondence:

$$\text{decay of } \hat{f} \leftrightarrow \text{"regularity"/"smoothness" of } f$$

Intuition - For $n$ large, $e_n$ is "squiggly". Too much squiggliness limits the smoothness of $f$.

Prop (Riemann-Lebesgue Lemma) - If $f \in L^1$, then $\hat{f} \in L^\infty$. Moreover, $\hat{f}$ vanishes at infinity. That is, $\lim_{|n|\to\infty} \hat{f}(n) = 0$.

Remark - This does not extend to finite measures. That is, $\mu$ a finite measure on $[0,1]$ does not imply $\hat{\mu}(n) \to 0$ at infinity. For example, let $\mu = \delta_0$, the delta measure centered at 0. Then for any $n$, we have $\hat{\mu}(n) = \int_0^1 e_n \, d\delta_0 = e_n(0) = 1$.

However, the Riemann-Lebesgue Lemma says that if $\mu \ll \lambda$, then $\hat{\mu}(n) \to 0$ at infinity.
Proof - The idea is to approximate $f$ by trigonometric polynomials. Let $f \in L^1$. Fix $\epsilon > 0$. Then there is $N$ such that $\|f - \sigma_N f\|_1 < \epsilon$. For $n > N$, we have

$$\hat{f}(n) = (\widehat{f - \sigma_N f})(n) + (\widehat{\sigma_N f})(n) = (\widehat{f - \sigma_N f})(n)$$

Thus, by Hölder's inequality, we have

$$|\hat{f}(n)| = |\int_0^1 (f - \sigma_N f)\overline{e_n}\, dx| \leqslant \|f - \sigma_N f\|_1 \|e_n\|_\infty < \epsilon$$

□

Remark - Previously we showed that $f \in L^2$ implies $\hat{f} \in \ell^2$. This can be seen as another regularity result.

**Differentiability**

Definition - We say $f \in L^2_{\text{per}}$ is *weakly differentiable* with *weak derivative* $g$ if, for all $\phi \in C^\infty_{\text{per}}$, we have

$$\int_0^1 f\phi' = -\int_0^1 g\phi$$

(Intuition: We require that integration by parts works.)
Example - If $f \in C^1_{\text{per}}$, then $f'$ is certainly a weak derivative of $f$ (as we would expect).
Proposition - If $f \in L^2_{\text{per}}$ and has weak derivative $f' \in L^2_{\text{per}}$, then

$$\hat{f}'(n) = 2\pi i n \hat{f}(n)$$

Remark - This is what we would hope for. If $f = \sum_n \hat{f}(n) e^{2\pi i n x}$, then differentiation term-by-term gives us the result above.

**Lecture 32 - 11/12/2014**

Last Time - Goal: "Regularity" of $f \leftrightarrow$ Decay of $\hat{f}$.
Proposition - If $f \in L^2_{\text{per}}[0,1]$, and has a weak derivative $f' \in L^2_{\text{per}}[0,1]$, then $\hat{f}'(n) = 2\pi i n \hat{f}(n)$.
Proof - By definition of a weak derivative,

$$\hat{f}'(n) = \langle f', e_n \rangle = -\langle f, e_n' \rangle = 2\pi i n \langle f, e_n \rangle = 2\pi i n \hat{f}(n)$$

$\square$

Corollary - If $f \in L^2_{\text{per}}$ has a weak derivative in $L^2_{\text{per}}$, then

$$\sum_n (1+|n|^2)|\hat{f}(n)|^2 < \infty$$

Proof - Since $f' \in L^2_{\text{per}}$, then $(\hat{f}'(n)) \in \ell^2$. Now use $\hat{f}'(n) = 2\pi i n \hat{f}(n)$. $\square$

Definition - Let $s \geqslant 0$. Define the (periodic) *Sobolev space of order s* by

$$\begin{aligned} H^s &:= \{f \in L^2_{\text{per}}[0,1] \mid \sum_n (1+|n|^2)^s |\hat{f}(n)|^2 < \infty\} \\ &= \{f \in L^2_{\text{per}}[0,1] \mid ((1+|n|^2)^{s/2}\hat{f}(n)) \in \ell^2\} \end{aligned}$$

For $f \in H^s$, denote $\|f\|_{H^s} = \|(1+|n|^2)^{s/2}\hat{f}(n)\|_{\ell^2}$.
Remark - For $s \geqslant 0$, $s \in \mathbb{N}$, we have

$$H^s = \{f \mid f \in L^2 \text{ and has } s^{\text{th}} \text{ order weak derivatives in } L^2\}$$

(If $s = 1$, note that $f \in H^1$ implies that $(2\pi i n \hat{f}(n)) \in \ell^2$, so this must correspond to some $g \in L^2_{\text{per}}$. By definition, this $g$ must be a weak derivative for $f$. For $s > 1$, this follows by induction.)

Theorem (Sobolev Embedding Theorem) - If $f \in H^s$ and $s > \frac{1}{2}$, then $f$ is continuous. That is, $f$ agrees a.e. with a continuous function.
Corollary - If $f \in H^s$, $s > n + \frac{1}{2}$, then $f \in C^n$.
Proof (of Theorem) - It's enough to prove the following result:

$$\exists C > 0,\ \|f\|_\infty \leqslant C\|f\|_{H^s}\ \forall f \in H^s \cap C^\infty_{\text{per}} \qquad (*)$$

That is, the inclusion map $H^s \cap C^\infty_{\text{per}} \hookrightarrow L^\infty$ is continuous.

Assuming $(*)$, let $f \in H^s$ and let $(\phi_n) \subseteq H^s \cap C^\infty_{\text{per}}$ such that $\phi_n \xrightarrow{H^s} f$ (e.g. $\phi_n = S_n f$). Now $(\phi_n)$ is Cauchy in $H^s$, so by $(*)$, $(\phi_n)$ is Cauchy in $L^\infty$, thus it converges uniformly. In particular, $\phi_n \to f$ uniformly, so $f$ is continuous.

It remains to prove $(*)$. For any $x \in [0,1]$, we have $f(x) = \sum_n \hat{f}(n)e_n(x)$. Therefore,

$$\|f\|_\infty \leqslant \sum_n |\hat{f}(n)| = \sum_n (1+|n|^2)^{s/2}|\hat{f}(n)| \cdot \tfrac{1}{(1+|n|^2)^{s/2}} \leqslant \|f\|_{H^s}\Big(\sum_n \tfrac{1}{(1+|n|^2)^s}\Big)^{1/2},$$

where the final step follows from the Cauchy-Schwarz inequality and the fact that $C := (\sum_n \frac{1}{(1+|n|^2)^s})^{1/2} < \infty$ for $s > \frac{1}{2}$. $\square$

Goal - Lebesgue differentiation

Suppose $\mu$ is a finite measure on $\mathbb{R}^d$ and $\mu \ll \lambda$. By Radon-Nikodym, there is $f \in L^1$ such that $d\mu = f\, d\lambda$. We *expect* that

$$f(x) = \lim_{r\to 0^+} \frac{\mu(B(x,r))}{\lambda(B(x,r))}$$

Lebesgue proved that this is true a.e.

Step 1: Considering the *maximal function*, defined by

$$M\mu(x) := \sup_{r>0} \frac{|\mu|(B(x,r))}{\lambda(B(x,r))}$$

If $f \in L^1$, then define $Mf(x) := \sup_{r>0} \frac{1}{\lambda(B(x,r))} \int_{B(x,r)} |f|\, d\lambda$. What can we prove about these objects?

Theorem - If $\mu$ is a finite measure, then $\lambda(\{M\mu > \alpha\}) \leqslant \frac{3^d}{\alpha}\|\mu\|$. If $f \in L^1$, then $\lambda(\{Mf > \alpha\}) \leqslant \frac{3^d}{\alpha}\|f\|_1$.

Want - To prove this theorem, we'd like $f \in L^1 \implies Mf \in L^1$. Unfortunately this is false, but we will get $Mf \in L^{1,\infty}$, which is enough to prove the theorem. Incidentally, we get something nicer for $L^p$, $p > 1$.

Fact (Homework) - If $f \in L^p$, $p \in (1,\infty]$, then $Mf \in L^p$ and $\|Mf\|_p \leqslant C_p\|f\|_p$, where $C_p$ depends only on $d$ and $p$, not on $f$.

**Lecture 33 - 11/14/2014**

Homework Discussion - Given $f \in C^1_{\text{per}}[0,1]$, it's easy to see that $\|f - \tau_h f\|_\infty \approx O(h)$, since $\|f - \tau_h f\| \leqslant \|f'\|_\infty h$. On your homework, you'll show something more general: If $f \in L^2$ and has weak derivative $Df \in L^2$, then $\|f - \tau_h f\|_2 \approx O(h)$.

Recall - Lebesgue Differentiation Theorem: Given $\mu$ a finite signed measure on $\mathbb{R}^d$, $\mu \ll \lambda$, we have $\frac{d\mu}{d\lambda}(x) = \lim_{r\to 0^+} \frac{\mu(B(x,r))}{\lambda(B(x,r))}$ holds $\lambda$-a.e. To prove this, we first introduce new quantities that will (eventually) provide upper bounds in our proof.

Definition - Let $\mu$ be a finite measure on $\mathbb{R}^d$, $f \in L^1(\mathbb{R})$. Define

$$M\mu(x) = \sup_{r>0} \frac{|\mu|(B(x,r))|}{\lambda(B(x,r))}$$

$$Mf(x) = \sup_{r>0} \frac{1}{\lambda(B(x,r))} \int_{B(x,r)} |f|$$

Theorem (Hardy-Littlewood Maximal Inequality) - There is a constant $C > 0$ (can choose $C = 3^d$) such that if $\mu$ is a finite measure on $\mathbb{R}^d$, then for all $\alpha > 0$ we have

$$\lambda(\{M\mu > \alpha\}) \leqslant \frac{C}{\alpha}\|\mu\|$$

Proof - IOU. $\square$

Remark - Given $f \in L^1$, we'd like to conclude that $Mf \in L^1$, perhaps $\|Mf\|_1 \leqslant C\|f\|_1$ for some constant $C$. This is false! For example, take $f = \chi_{[0,1]}$. For $x$ large, we have $Mf(x) \geqslant \frac{1}{2x} \notin L^1$!

Proof (of Lebegue Differentiation) - By Radon-Nikodym, it's enough to show that for any positive function $f \in L^1$, we have

$$\lim_{r\to 0^+} \frac{1}{\lambda(B(x,r))} \int_{B(x,r)} |f|\, d\lambda = f(x) \qquad \lambda\text{-a.e.}$$

The main idea is approximate $f$ by a continuous function, since the result clearly holds for continuous functions.

Fix $f \in L^1$. Define

$$\Omega f(x) = \limsup_{r>0} \tfrac{1}{\lambda(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, dy$$

To prove the result, it's enough to show that $\Omega f = 0$ a.e. (since $\int |f| \leqslant \int |f - f(x)| + f(x)$). Note that if $f$ is continuous, we are done. If not, fix $\epsilon > 0$ and write $f = g + h$, where $g \in L^1$ is continuous and $\|h\|_1 < \epsilon$. Note that

$$\Omega f \leqslant \Omega g + \Omega h = \Omega h$$

To estimate $\Omega h$, it's enough to consider the following (surprisingly loose) upper bound:

$$\begin{aligned} \Omega h(x) &= \limsup_{r>0} \tfrac{1}{\lambda(B(x,r))} \int_{B(x,r)} |h(y) - h(x)|\, dy \\ &\leqslant \left( \sup_{r>0} \tfrac{1}{\lambda(B(x,r))} \int_{B(x,r)} |h| \right) + |h(x)| \\ &= (Mh + |h|)(x) \end{aligned}$$

By the Hardy-Littlewood Maximal Inequality and Chebyshev's Inequality, we have, for $\alpha > 0$,

$$\begin{aligned} \lambda(\{\Omega f > \alpha\}) &\leqslant \lambda(\{\Omega h > \alpha\}) \\ &\leqslant \lambda(\{Mh > \tfrac{\alpha}{2}\}) + \lambda(\{|h| > \tfrac{\alpha}{2}\}) \\ &\leqslant \tfrac{3^d}{\alpha/2} \|h\|_1 + \tfrac{1}{\alpha/2} \|h\|_1 \\ &< \tfrac{2 \cdot 3^d + 2}{\alpha} \cdot \epsilon \end{aligned}$$

The quantity $\lambda(\{\Omega f > \alpha\})$ is independent of $\epsilon$, so take the limit $\epsilon \to 0^+$ to get $\lambda(\{\Omega f > \alpha\}) = 0$. Thus, the set $\{\Omega f > 0\}$ is $\lambda$-null, so $\Omega f = 0$ a.e. □

It remains to prove the Hardy-Littlewood Maximal Inequality. To do this, we state a lemma (which we'll prove next time).

Lemma (Vitali Covering Lemma) - Let $E \subseteq \mathbb{R}^d$, $\{B_1, \dots, B_N\}$ a set of balls such that $E \subseteq \bigcup_{1 \leqslant i \leqslant N} B_i$. Then there is a subcollection $\{B_i'\} \subseteq \{B_i\}$ of *disjoint* balls such that $E \subseteq \bigcup_i 3B_i'$.

Notation - If $B = B(x, r)$, then $3B := B(x, 3r)$.

Proof (Hardy-Littlewood) - It's enough to prove the result for $\mu$ positive (exercise). Let $\alpha > 0$ and define $E = \{M\mu > \alpha\}$. Fix $K \subseteq E$ compact. For each $x \in K$, there is $r_x > 0$ such that $\mu(B(x, r_x)) > \alpha\lambda(B(x,r))$. These balls cover $K$, so let $\{B_1, \dots, B_N\}$ be a finite subcover of $K$. By Vitali, there is a disjoint subcollection $\{B_1', \dots, B_M'\}$ such that $K \subseteq 3B_i'$. Now

$$\lambda(K) \leqslant \sum_{i=1}^{M} \lambda(3B_i) = 3^d \sum_{i=1}^{M} \lambda(B_i') < \tfrac{3^d}{\alpha} \sum_{i=1}^{M} \mu(B_i') = \tfrac{3^d}{\alpha} \mu(\bigcup_{i=1}^{M} B_i') \leqslant \tfrac{3^d}{\alpha} \|\mu\|$$

Taking the supremum over compact $K \subseteq E$, we have $\lambda(E) \leqslant \frac{3^d}{\alpha} \|\mu\|$, as desired. □

**Lecture 34 - 11/17/2014**

Lemma (Vitali) - Let $E \subseteq \mathbb{R}^d$, $\{B_1, \ldots, B_N\}$ a collection of balls whose union covers $E$. Then there is a disjoint subcollection $\{B_i'\}$ such that the balls $\{3B_i'\}$ cover $E$. (Here, $B = B(x,r) \implies 3B := B(x, 3r)$.)

Proof - By reordering, we may assume that radius$(B_1) \geqslant$ radius$(B_2) \geqslant \cdots$. Choose $B_1' = B_1$. Consider all balls disjoint from $B_1'$. Choose $B_2'$ to be the one of largest radius. Note that if $B_k \cap B_1' \neq \varnothing$ for some $k$, then $3B_1' \supseteq B_k$.

Given $\{B_1', \ldots, B_k'\}$, let $B_{k+1}'$ be the ball of largest radius that is disjoint from $B_1', \ldots, B_k'$ (if there is such a ball). Eventually, this process terminates. Now

- These balls are clearly disjoint.
- These balls cover $E$. Why? Given $B_i$, either $B_i = B_j'$ for some $j$, or $B_i$ meets some $B_j'$ of larger radius, in which case $B_i \subseteq 3B_j'$.

□

Corollary (of Lebesgue Differentiation) - Let $A \in \mathcal{L}(\mathbb{R}^d)$. Then

$$\lim_{r \to 0^+} \tfrac{\lambda(A \cap B(x,r))}{\lambda(B(x,r))} = \chi_A(x) \qquad \lambda\text{-a.e.}$$

Proof - Apply Lebesgue differentiation with $f = \chi_A$. □

Remark - Given $\mu$ a finite measure on $\mathbb{R}^d$, we know

$$\mu \ll \lambda \implies \lim_{r \to 0^+} \tfrac{\mu(B(x,r))}{\lambda(B(x,r))} = \tfrac{d\mu}{d\lambda}(x) \qquad \lambda\text{-a.e.}$$

What if $\mu \not\ll \lambda$? Decompose $\mu = \mu_S + \mu_{AC}$, where $\mu_{AC} \ll \lambda$ and $\mu_S \perp \lambda$. We know

$$\lim_{r \to 0^+} \tfrac{\mu_{AC}(B(x,r))}{\lambda(B(x,r))} = \tfrac{d\mu_{AC}}{d\lambda}(x) \qquad \lambda\text{-a.e.}$$

Claim (Homework) - $\lim_{r \to 0^+} \tfrac{|\mu_S|(B(x,r))}{\lambda(B(x,r))} = \begin{cases} 0 & \lambda\text{-a.e.} \\ \infty & \mu_S\text{-a.e.} \end{cases}$

* * *

**The Fundamental Theorem of Calculus**

Recall - For Riemann integration, if we define $F(x) = \int_0^x f \, dx$, where $f$ is continuous, then $F$ is differentiable and $F' = f$.

Theorem - Let $f \in L^1(\mathbb{R})$ and define $F(x) = \int_0^x f$. Then $F$ is differentiable and $F' = f$ a.e.

Proof - For $h > 0$, we have $\frac{F(x+h) - F(x-h)}{2h} = \frac{1}{2h} \int_{x-h}^{x+h} f = \frac{\mu(B(x,h))}{\lambda(B(x,h))}$, where $\mu(A) = \int_A f$. By the Lebesgue Differentiation Theorem, we have

$$f(x) = \tfrac{d\mu}{d\lambda}(x) = \lim_{h \to 0^+} \tfrac{\mu(B(x,h))}{\lambda(B(x,h))} = F'(x) \qquad \lambda\text{-a.e.}$$

□

Recall - For Riemann integration, if $f \in C^1$, then $\int_a^b f' = f(b) - f(a)$. Another way to express this is

$$f \in C^1 \implies \forall x \int_0^x f' = f(x) - f(0)$$

We would like to a version of this result for differentiable functions, dropping the $C^1$ assumption.

- **Attempt 1**: Is it enough to assume that $f$ is differentiable a.e.? The answer is *no*, since (for example), we may have $f' \notin L^1[0,1]$. For example, let $f(x) = \ln x$ for $x \in (0,1]$, $f(0) = 0$. Then $f$ is differentiable a.e. with $f' = \frac{1}{x}$, but $\int_0^x f' = \infty$ for all $x > 0$. In particular, $f' \notin L^1[0,1]$.
- **Attempt 2**: What if we assume $f$ is differentiable a.e. and $f' \in L^1$? This is still not enough. For example, let $f(x)$ be the Cantor function, which satisfies $f' = 0$ a.e. In particular, $\int_0^1 f' = \int_0^1 0 = 0 \neq 1 = f(1) - f(0)$.

To obtain the desired result, we introduce a new notion:

Definition - We say $f : [a,b] \to \mathbb{R}$ is *absolutely continuous* if, for any $\epsilon > 0$, there is $\delta > 0$ such that for all finite sets $\{(x_i, y_i)\}_{i \leqslant N}$ of disjoint subintervals of $[a,b]$, we have

$$\sum_{i=1}^{N} |x_i - y_i| < \delta \implies \sum_{i=1}^{N} |f(x_i) - f(y_i)| < \epsilon$$

Note - $f$ absolutely continuous implies $f$ continuous.

Proposition - Let $g \in L^1[a,b]$ such that $f(x) - f(a) = \int_a^x g$. Then $f$ is absolutely continuous.
Proof - Fix $\epsilon > 0$. We know $g$ is equi-integrable since $\lambda([a,b])$ is finite. Then there is $\delta > 0$ such that $\lambda(A) < \delta$ implies $\int_A |g| < \epsilon$. Now suppose that $\{(x_i, y_i)\}_{i \leqslant N}$ are disjoint with $\sum_{i=1}^{N} |x_i - y_i| < \delta$. Then $\lambda(\bigcup_{i=1}^{N} (x_i, y_i)) < \delta$, so $\int_{\bigcup_{i=1}^{N} (x_i, y_i)} |g| < \epsilon$. Now

$$\sum_{i=1}^{N} |f(x_i) - f(y_i)| = \sum_{i=1}^{N} \left| \int_{x_i}^{y_i} g \right| \leqslant \int_{\bigcup_{i=1}^{N} (x_i, y_i)} |g| < \epsilon$$

□

Theorem - A function $f : [a,b] \to \mathbb{R}$ is absolutely continuous if and only if

1. $f$ is differentiable a.e.
2. $f' \in L^1$
3. For all $x \in [a,b]$, $f(x) = f(a) + \int_a^x f'$

Proof - We just proved the reverse direction. We will prove the converse next time. □

**Lecture 35 - 11/19/2014**

Theorem - A function $f : [a, b] \to \mathbb{R}$ is absolutely continuous if and only if

1. $f$ is differentiable a.e.
2. $f' \in L^1$
3. For all $x \in [a, b]$, $f(x) = f(a) + \int_a^x f'$

Proof - We already proved the reverse direction last time. For the converse, suppose that $f$ is absolutely continuous.

**Lemma 1** Suppose $f$ is absolutely continuous and strictly increasing. Then (1), (2), and (3) hold.

**Remark** We need absolute continuity. Continuity + strictly increasing isn't enough, e.g. let $f(x) = g(x) + x$, where $g(x)$ is the Cantor function.

**Proof** Define $\mu(A) = \lambda(f(A))$. Note, $f^{-1}$ is increasing, thus measurable, so $A \in \mathcal{B}$ implies $f(A) \in \mathcal{B}$.

**Claim** $\mu \ll \lambda$

**Proof** Suppose $\lambda(A) = 0$. We need to show $\mu(A) = \lambda(f(A)) = 0$.

Fix $\epsilon > 0$ and let $K \subseteq f(A)$ be compact. We know $f^{-1}(K)$ is compact and $f^{-1}(K) \subseteq A$, so $\lambda(f^{-1}(K)) = 0$. Choose $\delta > 0$ such that, given disjoint $\{(x_i, y_i)\}_{i \leqslant N}$ with $\sum_{1 \leqslant i \leqslant N} |x_i - y_i| < \delta$¡ we have $\sum_{1 \leqslant i \leqslant N} |f(x_i) - f(y_i)| < \epsilon$. Since $\lambda(f^{-1}(K))$, we may choose such $\{(x_i, y_i)\}_{i \leqslant N}$ which cover $f^{-1}(K)$. Now

$$\lambda(K) \leqslant \sum_{1 \leqslant i \leqslant N} \lambda(f(x_i, y_i)) = \sum_{1 \leqslant i \leqslant N} |f(x_i) - f(y_i)| < \epsilon$$

Since $\epsilon > 0$ was arbitrary, then $\lambda(K) = 0$. Since $K \subseteq f(A)$ was arbitrary compact, then $\lambda(f(A)) = 0$. □

By Radon-Nikodym, there is $g \in L^1([a, b])$ such that $d\mu = g\, d\lambda$, so (1) and (2) hold with $f' = g$. For (3), write

$$f(x) - f(a) = \lambda((f(a), f(x))) = \mu((a, x)) = \int_a^x g\, d\lambda = \int_a^x f'\, d\lambda$$

□

**Lemma 2** Suppose $f$ is absolutely continuous and increasing. Then (1), (2), and (3) hold.

**Proof** Apply Lemma 1 to the function $g(x) = x + f(x)$. Now $g$ is differentiable a.e., $g' \in L^1$, so the same holds for $f$. For (3), write

$$\begin{aligned} f(x) = g(x) - x &= g(a) + \int_a^x g' - x \\ &= f(a) + a + \int_a^x (f' + 1) - x \\ &= f(a) + a + \int_a^x f' + (x - a) - x \\ &= f(a) + \int_a^x f' \end{aligned}$$

□

Definition - We say $f : [a, b] \to \mathbb{R}$ has *bounded variation* if the function

$$F(x) := \sup_{\Delta} \sum_{x_i \in \Delta} |f(x_i) - f(x_{i-1})|, \quad x \in [a, b]$$

is finite, where $\Delta$ ranges over all partitions of $[a, x]$. We call $F(x)$ the *variation* of $f$.

Remark - If $f : [a, b] \to \mathbb{R}$ has bounded variation, then $f$ can be written as a difference $g - h$ of increasing functions (in fact, the converse holds). This is because $f = \frac{(F+f)-(F-f)}{2}$, where $\frac{F \pm f}{2}$ are both increasing.

**Lemma 3** Suppose $f$ is absolutely continuous. Then (1), (2), and (3) hold.

**Proof** **Claim** If $f$ is absolutely continuous, then $f$ has bounded variation. Moreover, the variation of $f$ is absolutely continuous.

**Proof** Let $F$ denote the variation of $f$. Set $\epsilon = 1$ and let $\delta > 0$ such that if $\{(x_i, y_i)\}_{i \leqslant N}$ are disjoint and $\sum_{1 \leqslant i \leqslant N} |x_i - y_i| < \delta$, then $\sum_{1 \leqslant i \leqslant N} |f(x_i) - f(y_i)| < 1$. Let $\Delta$ be any partition of $[a, b]$. Refine it by setting

$$\Delta' = \Delta \cup (\{n\delta \mid n \in \mathbb{Z}\} \cap [a, b])$$

Now $\sum_{x_i \in \Delta} |f(x_i) - f(x_{i-1})| < \lceil \frac{b-a}{\delta} \rceil$. In particular, $F$ is finite. The proof that $F$ is absolutely continuous is left as an exercise. □

By the Claim, we may write $f = \frac{(F+f)-(F-f)}{2}$, a difference of increasing, absolutely continuous functions. By linearity of the integral and of differentation, we have the desired result. □

**Lecture 36 - 11/21/2014**

**Change of Variables**

Theorem - Let $U, V \subseteq \mathbb{R}^d$ be open, $\phi : U \to V$ a $C^1$ bijection. Then for all $f \in L^1(V)$, we have

$$\int_V f \, d\lambda = \int_U f \circ \phi |\det \nabla \phi| \, d\lambda,$$

where $\nabla \phi = (\partial_j \phi_i)$ is the Jacobian of $\phi$.
Proof - For $A \subseteq U$, $A \in \mathcal{B}(\mathbb{R}^d)$, define $\mu(A) = \lambda(\phi(A))$. *(Here, $\mu$ is defined so that $\lambda$ on $V$ is the pushforward of $\mu$ with respect to $\phi$.)* To prove the theorem, we need only prove a series of lemmas.

**Lemma 1** For all $A \in \mathcal{B}(U)$, $\phi(A) \in \mathcal{B}(V)$. Moreover, $\mu$ is a Borel measure on $U$.

**Lemma 2** $\mu \ll \lambda$

**Lemma 3** $\frac{d\mu}{d\lambda} = |\det \nabla \phi|$ $\lambda$-a.e.

Once these lemmas are proven, we have the result. This is because

$$\begin{aligned}\int_V f \, d\lambda &= \int_U f \circ \phi \, d\mu \quad \text{(pushforward)}\\ &= \int_U f \circ \phi \tfrac{d\mu}{d\lambda} \, d\lambda \quad (\mu \ll \lambda)\\ &= \int_U f \circ \phi |\det \nabla \phi| \, d\lambda\end{aligned}$$

**Proof (Lemma 1)** Suppose $A \in \mathcal{B}(U)$. Define

$$\Lambda = \{A \in U \mid \phi(A) \in \mathcal{B}(V)\}$$

This is a $\Lambda$-system because

- $\phi(U) = V \in \mathcal{B}(V)$
- If $A \subseteq B$ in $\Lambda$, then $\phi(B \backslash A) = \phi(B) \backslash \phi(A) \in \mathcal{B}(V)$ ($\phi$ is bijective).
- $(A_n) \subseteq \Lambda$ increasing implies $\phi(\bigcup_n A_n) = \bigcup_n \phi(A_n) \in \mathcal{B}(V)$.

Moreover, $\Lambda \supseteq \mathcal{K} := \{K \subseteq U \mid K \text{ is compact}\}$, which is a $\pi$-system ($\Lambda \supseteq \mathcal{K}$ because if $K \subseteq U$ is compact, then so is $\phi(K) \subseteq V$, so $\phi(K) \in \mathcal{B}(V)$). Thus, $\Lambda \supseteq \sigma(\mathcal{K}) = \mathcal{B}(U)$. (Compact sets generate the Borel $\sigma$-algebra since for any nonempty closed $C$ and $x \in C$, we have $C = \bigcup_n C \cap \overline{B(x,n)}$, a union of compact sets.) $\square$

**Proof (Lemma 3)** Assuming Lemma 2, we have $\frac{d\mu}{d\lambda}(x) = D\mu(x) := \lim_{r \to 0^+} \frac{\mu(B(x,r))}{\lambda(B(x,r))}$ holds $\lambda$-a.e., so it's enough to prove that $D\mu(x) = |\det \nabla \phi_x|$ for all $x \in U$. Below, we consider simple cases for $\phi$ and proceed to generalize.

- Step 0: Suppose $\phi$ is linear, i.e. $\phi(x) = Mx$ for some invertible $M \in M_d(\mathbb{R})$. Then $\mu = \lambda \circ \phi$ is translation invariant on $\mathbb{R}^d$ and finite on bounded sets, so by a previous result there is $C > 0$ such that $\mu(A) = C\lambda(A)$, so now $D\mu(x) = C$. Let $A$ be the unit cube. Then by a previous homework, we have

$$C = \mu(A) = |\det M| = |\det \nabla \phi| \implies D\mu(x) = |\det \nabla \phi_x|$$

- Step 1: We may assume without loss that $x = 0$ and $\phi(x) = 0$ (we can always reduce to this case by composing $\phi$ with translations).
  - **Case I**: Suppose $\det \nabla\phi_0 = 0$, i.e. $\nabla\phi_0$ is not invertible. We want to show $D\mu(0) = 0$. To do this, denote $T = \nabla\phi(0)$, $T : \mathbb{R}^d \to \mathbb{R}^d$. Since $\det T = 0$, then $\dim(\text{range}(T)) < d$, so $\lambda(\text{range}(T)) = 0$.
    Now let $\epsilon > 0$. By definition of differentiation, we have
    $$0 = \lim_{|x|\to 0} \tfrac{|\phi(x)-\phi(0)-\nabla\phi_0(x-0)|}{|x|} = \lim_{|x|\to 0} \tfrac{|\phi(x)-Tx|}{|x|}$$
    Then there is $r_0 > 0$ such that $|x| < r_0$ implies $|\phi(x) - Tx| < \epsilon|x|$. We know that for any $r$, $T(B(0,r))$ is contained in a $(d-1)$-dimensional subspace of $\mathbb{R}^d$ and $\text{diam}(T(B(0,r))) \leqslant \|T\| \cdot 2r$. Thus,
    $$\mu(B(0,r)) = \lambda(\phi(B(0,r))) \leqslant (\|T\|2r)^{d-1} \cdot \epsilon r$$
    Now
    $$0 \leqslant D\mu(0) \leqslant \lim_{r\to 0^+} \tfrac{(2\|T\|)^{d-1}\epsilon r^d}{r^d\lambda(B(0,1))} = (2\|T\|)^{d-1}\epsilon$$
    Taking the limit $\epsilon \to 0^+$, we have $D\mu(0) = 0 = \det \nabla\phi_0$
    *To be continued...*

**Lecture 37 - 11/24/2014**

Theorem - Let $U, V \subseteq \mathbb{R}^d$ be open, $\phi : U \to V$ a $C^1$ bijection. Then for all $f \in L^1(V)$, we have

$$\int_V f \, d\lambda = \int_U f \circ \phi |\det \nabla\phi| \, d\lambda,$$

where $\nabla\phi = (\partial_j \phi_i)$ is the Jacobian of $\phi$.
Proof (continued from last time) - We defined $\mu(A) = \lambda(\phi(A))$, a measure on $\mathcal{B}(U)$. We still need to show that $\mu \ll \lambda$ and $D\mu := \lim\limits_{r\to 0^+} \frac{\mu(B(\cdot,r))}{\lambda(B(\cdot,r))} = |\det \nabla\phi|$. We consider $D\mu(x)$ for $x = 0$, $\phi(x) = 0$.

- We've already proven that $D\mu = |\det \nabla\phi|$ if $\phi$ is linear or $\det \nabla\phi_0 = 0$.
- **Case II(a)**: Suppose $\nabla\phi_0 = I$, the identity matrix. Then $\lim\limits_{|x|\to 0} \frac{|\phi(x)-x|}{|x|} \to 0$, so there is $r_0 > 0$ such that $|\phi(x) - x| < \epsilon|x|$ for $|x| < r_0$. Now $\phi(B(0,r)) \subseteq B(0,(1+\epsilon)r)$ implies

$$\mu(B(0,r)) = \lambda(\phi(B(0,r))) \leqslant \lambda(B(0,(1+\epsilon)r)) = (1+\epsilon)^d \lambda(B(0,r))$$

In particular, $D\mu(0) = \lim\limits_{r\to 0^+} \frac{\mu(B(0,r))}{\lambda(B(0,r))} \leqslant (1+\epsilon)^d$. Taking the limit $\epsilon \to 0^+$, we have $D\mu(0) \leqslant 1$.

Remark - It turns out to be difficult to show that $D\mu \geqslant 1$. What we'd like is to show that $\phi(B(0,r)) \supseteq B(0,(1-\epsilon)r)$ (easy if $\phi^{-1} \in C^1$). This involves Brouwer's fixed-point theorem from topology, so we omit the proof here. (For more details, see pp. 150-151 of Rudin's Real and Complex Analysis, 3rd edition.) If you're unsatisfied, simply assume that $\phi^{-1} \in C^1$ and then the result follows (see Remark after this proof).

- **Case II(b)**: $\det \nabla\phi_0 \neq 0$. Set $T = \nabla\phi_0$ and consider $\psi = T^{-1}\phi$. Certainly $\nabla\psi_0 = I$, so by Case II(a),

$$\begin{aligned} 1 &= \lim_{r\to 0^+} \tfrac{\lambda(\psi(B(0,r)))}{\lambda(B(0,r))} \\ &= \lim_{r\to 0^+} \tfrac{\lambda(T^{-1}\phi(B(0,r)))}{\lambda(B(0,r))} \\ &= |\det T|^{-1} \lim_{r\to 0^+} \tfrac{\lambda(\phi(B(0,r)))}{\lambda(B(0,r))} \\ \implies |\det T| &= \lim_{r\to 0^+} \tfrac{\mu(B(0,r))}{\lambda(B(0,r))}, \end{aligned}$$

so we are done. This finishes the proof of Lemma 3 from last time. □

It remains to show that $\mu \ll \lambda$. By writing $U$ as an increasing union of compact sets, we have that $\mu$ is regular (Homework 3, Question 2). Thus, it suffices to show that if $K \subseteq U$ is compact and $\lambda(K) = 0$, then $\mu(K) = 0$.

Given such $K$, let $\epsilon > 0$ and let $W \supseteq K$ be open with $\lambda(W) < \epsilon$ and $C := \sup\limits_{x\in W} |\nabla\phi_x| < \infty$ (this is possible since we may choose $W$ to have compact closure). By the Mean Value Theorem, for any $x, y \in B(x_0, r_0) \subseteq W$, we have, for $i = 1, \ldots, d$,

$$\begin{aligned} \phi_i(x) - \phi_i(y) &= \nabla\phi_i(z_i) \cdot (x - y) \\ \implies |\phi_i(x) - \phi_i(y)| &\leqslant C|x - y| \\ \implies |\phi(x) - \phi(y)| &\leqslant dC|x - y| \\ \implies \lambda(\phi(B(x_0, r_0))) &\leqslant (dC)^d \lambda(B(x_0, r_0)) \end{aligned}$$

Now cover $K$ with balls $\{B(x,r_x) \mid x \in W, B(x,3r_x) \subseteq W\}$. By compactness, we may pass to a finite subcover. By the Vitali covering lemma, we may pass to a finite subset of disjoint balls $\{B_i'\}$ such that $K \subseteq \bigcup_i 3B_i'$. Now

$$\begin{aligned}
\mu(K) = \lambda(\phi(K)) &\leqslant \lambda(\bigcup_i \phi(3B_i')) \\
&\leqslant \sum_i \lambda(\phi(3B_i')) \\
&\leqslant \sum_i (dC)^d 3^d \lambda(B_i') \\
&= (dC)^d 3^d \lambda(\bigcup_i B_i') \\
&\leqslant (dC)^d 3^d \lambda(W) \\
&< (dC)^d 3^d \epsilon
\end{aligned}$$

Taking the limit $\epsilon \to 0^+$, we have $\mu(K) = 0$, so we are done. This finishes the proof of the theorem. □

Remark - Another way to show that the theorme holds for $\phi^{-1} \in C^1$ is through an appeal to symmetry. If we only prove that $D\mu \leqslant 1$ in Case II(a) above, we end up showing that

$$\int_V f \leqslant \int_U f \circ \phi |\det \nabla\phi| \, d\lambda \quad \text{for all } f \geqslant 0$$

Applying this result to $\phi^{-1}$, we have

$$\int_U f \circ \phi |\det \nabla\phi| \, d\lambda \leqslant \int_V f \circ \phi \circ \phi^{-1} |\det \nabla\phi_{\phi^{-1}(x)}| |\det \nabla\phi^{-1}| \, d\lambda = \int_V f \, d\lambda,$$

which implies equality.