Measure Theory

Lectures delivered by D. Stroock
Notes by Holden Lee

Spring 2011, MIT

Last updated Mon. 5/2/2011

Contents

Introduction

D. Stroock taught a course (18.125) on Measure Theory at MIT in Spring 2011. These are my "live-TeXed" notes from the course. The template is borrowed from Akhil Mathew.

Please email corrections to `holden1@mit.edu`.

Lecture 1
Wed. 2/2/2011

§1 Riemann integration

To integrate a function $f : J \to R$, where $J = [a_1, b_1] \times \cdots \times [a_N, b_N]$, take a non-overlapping cover $\mathcal{C}$ of J by nonoverlapping rectangles (i.e. for distinct $I, I' \in \mathcal{C}$, $I^\circ \cap I'^\circ = \phi$). Let $\xi \in \Xi(\mathcal{C})$ be a choice function that assigns to each I an element in I ($\xi(I) \in I$). Let

$$\mathcal{R}(f; \mathcal{C}, \xi) = \sum_{I \in \mathcal{C}} f(\xi(I))\mathrm{vol}(I)$$

where $\mathrm{vol}(I)$ is the product of its sides.

One says that f is Riemann integrable if there exists $A \in \mathbb{R}$ such that for all $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|\mathcal{R}(f, \mathcal{C}, \xi) - A| < \epsilon \text{ for all } \mathcal{C} \text{ with } ||\mathcal{C}|| < \delta, \xi \in \Xi(\mathcal{C}).$$

where

$$||\mathcal{C}|| = \max_I \mathrm{diam}(I).$$

This value of A is denoted by

$$A = (R) \int_J f(x)\, dx$$

Theorem 1.1: Any continuous function is Riemann integrable.

Proof. Uniform continuity of f (from compactness of domain) gives that approximations get close; completeness of $\mathbb{R}$ gives existence of A. □

Lemma 1.2: Suppose that $\mathcal{C}$ is any collection of rectangles I.

1. If $\mathcal{C}$ is non-overlapping and $J \supseteq \bigcup \mathcal{C}$, then $\mathrm{vol}(J) \geq \sum_{I \in \mathcal{C}} \mathrm{vol}(I)$.
2. If $J \subseteq \bigcup \mathcal{C}$, then $\mathrm{vol}(J) \leq \sum_{I \in \mathcal{C}} \mathrm{vol}(I)$.

Proof. Without loss of generality, we may assume $J \subseteq \bigcup \mathcal{C}$ (just intersect rectangles with J), and $I^\circ \neq \phi$ for any $I \in \mathcal{C}$.

Induct on number of dimensions N. Consider $N = 1$. Let $I = [a_I, b_I]$.

For the first part, choose $a_J \leq c_0 < \cdots < c_l \leq b_J$ such that

$$\{c_k : 0 \leq k \leq l\} = \{a_I : I \in \mathcal{C}\} \cup \{b_I : I \in \mathcal{C}\}.$$

Let $\mathcal{C}_k = \{I \in \mathcal{C} : [c_{k-1}, c_k] \subseteq I\}$. Note

1. $\mathrm{vol}(I) = \sum_{k, I \in \mathcal{C}_k} (c_k - c_{k-1}) = b_I - a_I$.

2. If $\mathcal{C}$ is non-overlapping, then I is in at most one $\mathcal{C}_k$ (by definition of c_i as endpoints).

Then

$$\sum_{I\in\mathcal{C}} \mathrm{vol}(I) = \sum_{I\in\mathcal{C}} \sum_{k:I\in\mathcal{C}_k} (c_k - c_{k-1}) \le \sum_{k=1}^{l} \sum_{I\in\mathcal{C}_k} (c_k - c_{k-1}) \le c_l - c_0 \le b_J - a_J = \mathrm{vol}(J).$$

For the second part, if $J = \bigcup \mathcal{C}$ then $c_0 = a_J$, $c_l = b_J$, and $\mathcal{C}_k \ne \phi$ for any $1 \le k \le l$. (For this second assertion, consider 2 cases: c_k is the left or right hand endpoint. Argument is the same. For the right endpoint, choose I so that $b_I \ge c_k$ and $a_I \le a_{I'}$ for every I' such that $b_{I'} \ge c_k$—i.e. left-hand endpoint is as small as possible. Then $a_I \le c_{k-1}$; else any interval starting at c_{k-1} ends before c_k, contradiction.) Now

$$\sum_{I\in\mathcal{C}} \mathrm{vol}(I) = \sum_{I\in\mathcal{C}} \sum_{k:I\in\mathcal{C}_k} (c_k - c_{k-1}) \ge \sum_{k=1}^{l} (c_k - c_{k-1}) = b_J - a_J.$$

When $N > 1$, we can write $I = R_I \times [a_I, b_I]$ where R_I is a $(n-1)$-dimensional rectangle. Apply a similar argument, but with $R_J = \bigcup_{I\in\mathcal{C}_k} R_I$. □

To "remove" the choice function we consider the Riemann upper and lower sums.

$$\mathcal{U}(f;\mathcal{C}) = \sum_{I\in\mathcal{C}} (\sup_I f)\,\mathrm{vol}(I) \ge \mathcal{R}(f;\mathcal{C},\xi)$$
$$\mathcal{L}(f;\mathcal{C}) = \sum_{I\in\mathcal{C}} (\inf_I f)\,\mathrm{vol}(I) \le \mathcal{R}(f;\mathcal{C},\xi)$$

Proposition 1.3: Let $f : J \to \mathbb{R}$ be bounded. f is Riemann integrable if and only if

$$\lim_{||\mathcal{C}||\to 0} \mathcal{L}(f;\mathcal{C}) = \lim_{||\mathcal{C}||\to 0} \mathcal{U}(f;\mathcal{C}).$$

Proof. "⇐"—squeeze theorem. "⇒"—choose choice function so close to upper/lower sum. □

The lemma applies when $\mathcal{C}_2$ is a refinement of $\mathcal{C}_1$, written $\mathcal{C}_1 \le \mathcal{C}_2$ (every rectangle of $\mathcal{C}_2$ is inside a rectangle in $\mathcal{C}_1$). Then $\mathcal{U}(f;\mathcal{C}_1) \ge \mathcal{U}(f;\mathcal{C}_2)$ and $\mathcal{L}(f;\mathcal{C}_1) \le \mathcal{L}(f;\mathcal{C}_2)$. Since I_1 is covered by nonoverlapping intervals of $\mathcal{C}_2$; $\mathrm{vol}(I_1)$ is sum of volumes of those intervals.

Lecture 2
Fri. 2/4/2011

§1 Riemann integrability

Theorem 2.1: Let $f : J \to \mathbb{R}$ be bounded. Then

1. $\lim_{||\mathcal{C}||\to 0} \mathcal{U}(f, \mathcal{C}) = \inf_{\mathcal{C}} \mathcal{U}(f, \mathcal{C})$.
2. $\lim_{||\mathcal{C}||\to 0} \mathcal{L}(f, \mathcal{C}) = \sup_{\mathcal{C}} \mathcal{L}(f, \mathcal{C})$.
3. f is Riemann integrable if and only if
$$\inf_{\mathcal{C}} \mathcal{U}(f, \mathcal{C}) = \sup_{\mathcal{C}} \mathcal{L}(f, \mathcal{C}).$$
where the infimum and supremum are taken over all finite exact nonoverlapping coverings.

Proof. We use the following.

Lemma 2.2: Given $\mathcal{C}$ and $\varepsilon > 0$ there exists δ such that $||\mathcal{C}'|| \leq \delta$ such that $\mathcal{U}(f, \mathcal{C}') \leq U(f, \mathcal{C}) + \varepsilon$. (Note $\mathcal{C}'$ need not be a refinement.)

Similarly, there exists δ such that $||\mathcal{C}'|| \leq \delta$ such that $\mathcal{L}(f, \mathcal{C}') \geq L(f, \mathcal{C}) - \varepsilon$. (Note $\mathcal{C}'$ need not be a refinement.)

Proof. Consider $I' \in \mathcal{C}'$. Then either

1. $I' \subseteq I$ for $I \in \mathcal{C}$ (the "good" type) or
2. I' hits an edge (the "bad" case).

The terms in the first case do not cause a problem—if every I' were of this type then $\mathcal{U}(f, \mathcal{C}') \leq \mathcal{U}(f, \mathcal{C}')$.

The rectangles in the second case cannot have a large combined area for $||\mathcal{C}'||$ small—they must be in a δ-neighborhood of the edges. In fact

$$\left| \sum_{I'} (\sup_{I'} f) \operatorname{vol}(I') \right| \leq 2\delta ||f||_u C$$

where C depends on N, the cardinality of $\mathcal{C}$, and J, and the uniform norm is

$$||f||_u = \sup_{x \in J} |f(x)|.$$

$\square$

Choose $\mathcal{C}$ so the upper sum is close to the infimum:

$$\mathcal{U}(f, \mathcal{C}) \leq \inf_{\mathcal{C}} \mathcal{U}(f, \mathcal{C}) + \frac{\varepsilon}{2}.$$

Then find δ as in the lemma (for $\frac{\varepsilon}{2}$); for $||\mathcal{C}|| < \delta$, we have

$$\mathcal{U}(f, \mathcal{C}) \leq \inf_{\mathcal{C}} \mathcal{U}(f, \mathcal{C}) + \varepsilon.$$

Item 2 follows similarly.

Use Proposition 1.3 to get item 3. $\square$

§2 Riemann-Stieltjes integral

In the Riemann integral we integrate with respect to "homogeneous density", dx means summing $b_I - a_I$. For the Riemann integral we replace dx with $d\psi$, and sum $\psi(b_I) - \psi(a_I)$ instead of $b_I - a_I$.

Definition 2.3: The Riemann sum of ϕ over $\mathcal{C}$ with respect to ψ relative to ξ is

$$\mathcal{R}(\varphi|_\psi, \mathcal{C}, \xi) = \sum_{I \in \mathcal{C}} \varphi(\xi(I))\Delta_I \psi, \quad \Delta_I \psi = \psi(b_I) - \psi(a_I).$$

ϕ is Riemann integrable with respect to ψ if there exists $A \in \mathbb{R}$ such that for every $\varepsilon > 0$, there exists $\delta > 0$ such that for every $||\mathcal{C}|| \leq \delta$ and any ξ,

$$|\mathcal{R}(\varphi|\psi, \mathcal{C}, \xi) - A| < \varepsilon.$$

Then we write

$$(R) \int_J \varphi(x) d\psi(x) = A.$$

Proposition 2.4: For $\varphi \in \mathcal{C}(J, \mathbb{R}), \psi \in C^1(J, \mathbb{R})$,

$$(R) \int_J \phi(x)\, d\psi(x) = (R) \int_J \varphi(x)\psi'(x)\, dx.$$

Proof. By the Mean Value Theorem,

$$\psi(b_I) - \psi(a_I) = \psi'(\eta(I)) \operatorname{vol}(I).$$

Now use uniform continuity of ψ'. □

Example 2.5: Suppose $a = a_0 < a_1 < \ldots < a_n = b$, and ψ is constant on (a_{m-1}, a_m) for $m = 1, \ldots, n$ (a "step" function with a few naughty points), and $\varphi \in C(J; \mathbb{R})$. Then

$$(R) \int_J \varphi\, d\psi = \sum_{m=1}^{n-1} \varphi(a_m)(\psi(a_m+) - \psi(a_m-)) + \varphi(a)(\psi(a+) - \psi(a)) + \varphi(b)(\psi(b) - \psi(b-)).$$

Proof. Consider an interval $I \in \mathcal{C}$. We may assume $\mathcal{C}$ has a fine enough mesh so no interval contains more than one a_i. Then either

1. $I \cap \{a_0, \ldots, a_n\}$, i.e. $I \subseteq (a_{m-1}, a_m)$ for some m. ("Good case") Then $\Delta_I \psi = 0$.

2. $a_m \in I^\circ$. Then $\Delta_I \psi = \psi(a_m+) - \psi(a_m-)$.

3. $a_m = a_I$ or b_I. Then

$$\Delta_I \psi = \begin{cases} \psi(a_m+) - \psi(a_m), & a_m = a_I \\ \psi(a_m) - \psi(a_m-), & a_m = b_I \end{cases}.$$

□

Example 2.6:
$$(R)\int_J (\alpha\varphi_1+\beta\varphi_2)\,d\psi=\alpha\int_J \phi_1\,d\psi+\beta\int_J \varphi_2\,d\psi.$$
Example 2.7: Let $J=J_1\cup J_2$ and $J_1^\circ\cap J_2^\circ=\varphi$. Then
$$(R)\int_J \varphi\,d\psi=(R)\int_{J_1}\varphi\,d\psi+(R)\int_{J_2}\varphi\,d\psi.$$

Proof. We want
$$|\mathcal{R}(\varphi|\psi,\mathcal{C}_1,\xi_1)-\mathcal{R}(\varphi|\psi,\mathcal{C}_1',\xi_1')|<\varepsilon$$
Let $\mathcal{C}=\mathcal{C}_1\cup\mathcal{C}_2$ and $\xi=\xi_1\cup\xi_2$, and similarly with $\mathcal{C}'$ and ξ'. The difference above equals
$$|\mathcal{R}(\varphi|\psi,\mathcal{C},\xi)-\mathcal{R}(\varphi|\psi,\mathcal{C}',\xi')|$$
as needed.

Choose $\mathcal{C}_1,\mathcal{C}_2$ so the Riemann sums of the RHS integrals are close to the integral; then glue as above. □

Lecture 3
Mon. 2/7/2011

§1 Integration by parts

Theorem 3.1 (Integration by parts)**:** Suppose φ and ψ are bounded functions on J. If φ is Riemann integrable, then ψ is φ-Riemann integrable, and
$$(R)\int_J \psi(x)\,d\xi(x)=[\varphi(x)\psi(x)]_a^b-(R)\int_J\varphi(x)\,d\psi(x).$$

Proof. Let $\mathcal{C}=\{[\alpha_{m-1},\alpha_m]:1\le leqn\}$ where $a=\alpha_0<\ldots<\alpha_n=b$. Take $\varphi([\alpha_{m-1},\alpha_m])=\beta_m\in[\alpha_{m-1},\alpha_m]$. Now
$$\begin{aligned}\mathcal{R}(\varphi|\psi;\mathcal{C},\varphi)&=\sum_{m=1}^n\psi(\beta_m)(\varphi(\alpha_m)-\varphi(\alpha_{m-1}))\\&=\sum_{m=1}^n\psi(\beta_m)\varphi(\alpha_m)-\sum_{m=0}^{n-1}\psi(\beta_{m+1})\varphi(\alpha_m)\\&=\psi(\beta_n)\varphi(b)-\psi(\beta_1)\varphi(a)-\sum_{m=1}^{n-1}\varphi(\alpha_m)(\psi(\beta_{m+1})-\psi(\beta_m))\end{aligned}$$

Now we think of the β_m as the endpoints of intervals and α_m as the choices. Let $\beta_0 = a$ and $\beta_{n+1} = b$. Rearranging gives

$$\psi(b)\varphi(b) - \psi(a)\varphi(a) - \sum_{m=0}^{n} \varphi(\alpha_m)(\psi(\beta_{m+1}) - \psi(\beta_m)).$$

The mesh size of the β-partition is at most twice the mesh size of the α-partition. Since φ is ψ-integrable, this last sum approaches $(R)\int_J \varphi(x)\, d\psi(x)$, as needed. □

Corollary 3.2 (Fundamental Theorem of Calculus)**:** Suppose φ is differentiable. Then

$$\int_J \frac{d\xi(x)}{dx}\, dx = [\varphi(x)]_a^b$$

Proof. Take $\psi = 1$ and note $\int_J \frac{d\xi(x)}{dx}\, dx = \int_J d\xi(x)$. □

§2 Riemann-Stieltjes integrability

Theorem 3.3: If ψ is increasing, then every $\varphi \in C(J; \mathbb{R})$ is ψ-Riemann integrable.

Proof. Define $\mathcal{U}(\varphi|\psi; \mathcal{C})$ and $\mathcal{L}(\varphi|\psi; \mathcal{C})$. Use uniform continuity. □

Proposition 3.4: If ψ_1, ψ_2 are increasing, then for every $\varphi \in C(J; R)$,

$$(R)\int_J \varphi(x)\, d(\psi_1 - \psi_2) = (R)\int_J \varphi(x)\, d\psi_1 - (R)\int_J \varphi(x)\, d\psi_2.$$

§3 Variation

Proposition 3.5: If $\psi = \psi_2 - \psi_1$,

$$\left|(R)\int_J \varphi(x)\, d\psi(x)\right| \le ||\varphi||_u(\Delta_J\psi_1 + \Delta_J\psi_2).$$

We can ask the following: for what functions ψ does there exist K_ψ such that for every $\varphi \in C(J; R)$ ψ-integrable and

$$\left|(R)\int \varphi\, d\psi\right| \le K_\psi ||\varphi||_u?$$

Theorem 3.6: Only functions that are the difference of two increasing functions.

We give a better description of this criterion.

Definition 3.7: Let

$$S(\psi, \mathcal{C}) = \sum_{I \in \mathcal{C}} |\Delta_I \psi|.$$

The **variation** of ψ on J is

$$\mathrm{Var}(\psi, J) = \sup_{\mathcal{C}} S(\psi; \mathcal{C}).$$

This measures the amount of "up-and-down" jiggliness of the function.

Proposition 3.8 (Basic properties)**:** 1. By the Triangle Inequality, if $\mathcal{C}'$ is a refinement of $\mathcal{C}$, then $S(\psi;\mathcal{C}') \geq S(\psi;\mathcal{C})$.

2. If $J = J_1 \cup J_2$, $\mathrm{Var}(\psi, J) = \mathrm{Var}(\psi, J_1) + \mathrm{Var}(\psi, J_2)$.

Definition 3.9: For $a \in \mathbb{R}$ let $a_+ = \max(a, 0)$ and $a_- = \max(-a, 0)$. Define

$$\begin{aligned} S_+(\psi;\mathcal{C}) &= \sum_{I\in\mathcal{C}} (\Delta_I \psi)^+ \\ S_-(\psi;\mathcal{C}) &= \sum_{I\in\mathcal{C}} (\Delta_I \psi)^- \\ \mathrm{Var}_\pm(\psi;\mathcal{C}) &= \sup_{\mathcal{C}} S_\pm(\psi;\mathcal{C}). \end{aligned}$$

Note $a_+ - a_- = a$ and $a_+ + a_- = |a|$, so

$$\begin{aligned} S_+(\psi,\mathcal{C}) - S_-(\psi,\mathcal{C}) &= \Delta_J \psi \\ S_+(\psi,\mathcal{C}) + S_-(\psi,\mathcal{C}) &= S(\psi;\mathcal{C}) \\ S_+(\psi,\mathcal{C}) &= \frac{1}{2}(\Delta_J \psi - S(\psi;\mathcal{C})) \\ S_-(\psi,\mathcal{C}) &= \frac{1}{2}(\Delta_J \psi + S(\psi;\mathcal{C})) \end{aligned}$$

If approach extreme values for one than for all of them. Statements pass to variations.

$$\begin{aligned} \mathrm{Var}_+(\psi;J) - \mathrm{Var}_-(\psi;J) &= \Delta_J \psi \\ \mathrm{Var}_+(\psi,J) + \mathrm{Var}_-(\psi;J) &= \mathrm{Var}(\psi;J) \end{aligned}$$

Lecture 4
Wed. 2/9/2011

§1 Functions of Bounded Variation

Theorem 4.1: A function can be written as the difference of two increasing functions if and only if it has bounded variation.

Proof. Suppose that

$$\mathrm{Var}(\psi;[a,b]) < \infty.$$

Then

$$\begin{aligned} \psi(x) - \psi(a) &= V_+(\psi;[a,x]) - V_-(\psi;[a,x]) \\ \psi(x) &= [\psi(a) + V_+(\psi;[a,x])] - V_-(\psi;[a,x]) \end{aligned}$$

□

§2 Convergence rate

Let $f : [0,1] \to \mathbb{R}$ be smooth ($f \in C^1$) Let

$$R_n(f) = \frac{1}{n}\sum_{m=1}^{n} f\left(\frac{m}{n}\right).$$

$$(R)\int_0^1 f(x)\,dx = \lim_{n\to\infty} R_n(f).$$

How fast does $R_n(f)$ converge to the integral? In general we can't do better than the following argument:

$$\begin{aligned}
(R)\int_0^1 f(x)\,dx - R_n(f) &= \sum_{m=1}^{n}\int_{\frac{m-1}{n}}^{\frac{m}{n}}\left(f(x) - f\left(\frac{m}{n}\right)\right)\\
&= \int_{\frac{m-1}{n}}\left(x - \frac{m}{n}\right) f'(\xi_x)\,dx\\
&\le n \cdot ||f'||_u \frac{1}{n^2} = \frac{1}{n}||f'||_u
\end{aligned}$$

This is the best we can do because letting $f(x) = x$,

$$R_n(f) = \frac{1}{n}\sum_{m=1}^{n}\frac{m}{n} = \frac{n(n+1)}{2n^2} = \frac{1}{2} + \frac{1}{2n}.$$

However, if f is periodic we can do a lot better. We integrate by parts, in such a way so that we don't boundary terms, by choosing constants of integration appropriately. Then

$$\begin{aligned}
&(R)\int_0^1 f(x)\,dx - R_n(f)\\
&= -\sum_{m=1}^{n}(R)\int_{\frac{m-1}{n}}^{\frac{m}{n}}\left(f(x) - f\left(\frac{m-1}{n}\right)\right)\,dx\\
&= -\sum_{m=1}^{n}(R)\int_{\frac{m-1}{n}}^{\frac{m}{n}}\left(x - \frac{m-1}{n}\right) f'(x)\,dx\\
&= -\sum_{m=1}^{n}(R)\int_{\frac{m-1}{n}}^{\frac{m}{n}}\left(x - \frac{m-1}{n} - \frac{1}{2n}\right) f'(x)\,dx && \text{by periodicity}\\
&= -\sum_{m=1}^{n}(R)\int_{\frac{m-1}{n}}^{\frac{m}{n}}\left(x - \frac{m-1}{n} - \frac{1}{2n}\right)\left(f'(x) - f\left(\frac{m}{n}\right)\right)\,dx.
\end{aligned}$$

We used periodicity to subtract off the average value of $x - \frac{m-1}{n}$ on $\left[\frac{m-1}{n}, \frac{m}{n}\right]$. (Note $\int_0^1 cf'(x)\,dx = 0$.) In the last step we used $\frac{1}{2n}$ is the average value of $x - \frac{m-1}{n}$ on the integral, so $\int_{\frac{m-1}{n}}^{\frac{m}{n}} \left(x - \frac{m-1}{n} - \frac{1}{2n}\right) = 0$. The last expression is bounded by $\frac{1}{2n^2}||f''||_u$, better than $\frac{1}{n}$.

Now assume that f and f' are both periodic C^1. (i.e. $f'(1) = f'(0)$ as well)

$$\begin{aligned}
&\sum_{m=1}^{n} \int_{\frac{m-1}{n}}^{\frac{m}{n}} \left(x - \frac{m-1}{n} - \frac{1}{2n}\right) \left(f'(x) - f\left(\frac{m}{n}\right)\right) dx \\
&= -\frac{1}{2n} \sum_{m=1}^{n} \int_{\frac{m-1}{n}}^{\frac{m}{n}} f'(x) - f'\left(\frac{m}{n}\right) + \sum_{m=1}^{n} \int_{\frac{m-1}{n}}^{\frac{m}{n}} \left(x - \frac{m-1}{n}\right) \left(f'(x) - f'\left(\frac{m}{n}\right)\right) \\
&= -\frac{1}{2n} \sum_{m=1}^{n} \int_{\frac{m-1}{n}}^{\frac{m}{n}} f'(x) - f'\left(\frac{m}{n}\right) + \sum_{m=1}^{n} \int_{\frac{m-1}{n}}^{\frac{m}{n}} \left(x - \frac{m-1}{n}\right)^2 f''(x)
\end{aligned}$$

Use periodicity to subtract off average value.

$$-\frac{1}{2n} \sum_{m=1}^{n} \int_{\frac{m-1}{n}}^{\frac{m}{n}} f'(x) - f'\left(\frac{m}{n}\right) + \sum_{m=1}^{n} \int_{\frac{m-1}{n}}^{\frac{m}{n}} \left[\left(x - \frac{m-1}{n}\right)^2 - \frac{1}{3n^2}\right] \left(f''(x) - f''\left(\frac{m}{n}\right)\right) dx.$$

Repeating this process. If f and all its derivatives are periodic (of period 1) then the error in the Riemann approximation is going to 0 faster than any $\frac{1}{n^k}$.

To define Riemann integals for complex valued-functions, just look at real and complex parts separately. Let $f(x) = e^{2\pi i x}$. Let $\xi_{m,n} = \frac{m}{n}\left(1 - \frac{1}{n}\right)$. Then

$$n\left(\int_0^1 e^{2\pi i x}\, dx - \frac{1}{n}\sum_{m=1}^{n} f(\xi_{m,n})\right) = \frac{1}{n}\sum_{m=1}^{n} e^{\frac{2\pi i(1-\frac{1}{n})m}{n}} = e^{\frac{2\pi i(1-\frac{1}{n})}{n}} \frac{1 - e^{-\frac{2\pi i}{n}}}{1 - e^{\frac{2\pi i(1-\frac{1}{n})}{n}}} \to -1.$$

Moral: Periodicity destroyed by bad choice function.

Definition 4.2: The **Bernoulli numbers** b_l are inductively defined by

$$b_{l+1} = \sum_{k=0}^{l} \frac{(-1)^k}{(k+2)!} b_{l-k}.$$

§3 Lebesgue integration: Motivation

Let E be a set. Let $f : E \to \mathbb{R}$. Let μ be "volume" or measure of Γ; that is μ is defined on nice subsets of E. We want to find a way to integrate f with respect to μ. We want to find a partition of E into subsets Γ so that f is constant or close to constant on each set Γ. We then add up $\mu(E)y$ (y this constant value).

Riemann: E has topological structure and f is nice with respect to topology (e.g. continuous). Partition into sets small from the topological standpoint, then give me f, it'll be nearly constant on each subset.

But if f doesn't respect topology this FAILS!

Lebesgue: Look at sets of the form

$$\{x|f(x) \in [(m-1)2^{-n}, m2^{-n}]\}$$

f is nearly constant on each of these subsets, regardless of topological niceness. Now integrate.

But first we need to find the "volume" or "measure" of these sets! They will be HIDEOUS... Integration theory is easy compared to assigning measures.

Lecture 5
Fri. 2/11/2011

§1 Measure

For a set $E \neq \phi$ define the power set

$$\mathcal{P}(E) = 2^E = \{\Gamma : \Gamma \subseteq E\}.$$

Definition 5.1: A subset $\mathcal{B} \subseteq \mathcal{P}(E)$ is a **σ-algebra** it satisfies the following properties:

1. $E \in \mathcal{B}$.
2. $\mathcal{B}$ is closed under complementation: $\Gamma \in \mathcal{B}$ implies $\Gamma^c = E \backslash \Gamma \in \mathcal{B}$.
3. $\{\Gamma_n : n \geq 1\} \subseteq \mathcal{B}$ implies $\bigcup_{n=1}^{\infty} \Gamma_n \in \mathcal{B}$.

(If item 2 is satisfied just for finite instead of countable unions then we call $\mathcal{B}$ an algebra.)

Note that items 2 and 3 imply that a countable intersection of elements in $\mathcal{B}$ is in $\mathcal{B}$, and a difference of sets in $\mathcal{B}$ is in $\mathcal{B}$.

Definition 5.2: We call $(E, \mathcal{B})$ is a measurable space. A measure on $(E, \mathcal{B})$ is a map $\mu : \mathcal{B} \to [0, \infty]$ such that

1. $\mu(\phi) = 0$.
2. (Countable additivity) If $\{\Gamma_n : n \geq 1\}$ is a family of pairwise disjoint subsets of E, then
$$\mu\left(\bigcup_{n=1}^{\infty} \Gamma_n\right) = \sum_{n=1}^{\infty} \mu(\Gamma_n),$$
i.e. the volume of the whole is the sum of the volume of the parts.

Compare this to the definition of a topological space—measurable spaces have measureable sets while topologies have open sets.

Example 5.3: Define a measure μ on the integers $\mathbb{Z}$ by associating some $\mu_i \geq 0$ for each integer i, and setting

$$\mu(\Gamma) = \sum_{i \in \Gamma} \mu_i.$$

Our strategy is to start with some class of nice, well-defined subsets, and generate more.

Definition 5.4: For a family of subsets $\mathcal{C} \subseteq \mathcal{P}(E)$, define the σ-algebra generated by $\mathcal{C}$, denoted by $\sigma(\mathcal{C})$, to be the smallest σ-algebra containing $\mathcal{C}$. In other words it is the intersection of all σ-algebras containing $\mathcal{C}$. (This is well-defined since the power set is a σ-algebra containing $\mathcal{C}$.)

If E is a topological space and $\mathcal{C} = \{\Gamma \subseteq E : \Gamma \text{ open}\}$ then $\sigma(\mathcal{C}) = \mathcal{B}_E$ is called the **Borel σ-algebra**.

Lebesgue showed that there exists a unique Borel **σ-algebra** on $\mathcal{B}_{\mathbb{R}_N}$ such that $\mu_{\mathbb{R}^N}(I) = \mathrm{vol}(I)$.

§2 Basic results

Proposition 5.5:

1. If $A \subseteq B$ are sets in $\mathcal{B}$ then $\mu(A) \leq \mu(B)$.
2. (Countable subadditivity) Let $\{\Gamma_n : n \geq 1\} \subseteq \mathcal{B}$. Then

$$\mu\left(\bigcup_{n=1}^{\infty} \Gamma_n\right) \leq \sum_{n=1}^{\infty} \mu(\Gamma_n).$$

   (The sets are not necessarily disjoint, so the RHS counts "overlap.")

3. A countable union of subsets of measure zero has measure 0.
4. We write $B_n \nearrow B$ if $B_1 \subseteq B_2 \subseteq \cdots$ and $\bigcup_{n=1}^{\infty} B_n = B$.
   If $B_n \nearrow B$ then $\mu(B_n) \nearrow \mu(B)$ (i.e. $\mu(B_n) \to \mu(B)$ from below).
5. We write $B_n \searrow B$ if $B_1 \supseteq B_2 \supseteq \cdots$ and $\bigcap_{n=1}^{\infty} B_n = B$.
   If $B_n \searrow B$ *and* $\mu(B_1) < \infty$ then $\mu(B_n) \searrow \mu(B)$ (i.e. $\mu(B_n) \to \mu(B)$ from above).

Proof.

1. Note $B \backslash A \in \mathcal{B}$. Hence

$$\mu(B) = \mu(A) + \mu(B \backslash A) \geq \mu(A).$$

2. Let

$$B_n = \Gamma_n \backslash \bigcup_{m=1}^{n-1} \Gamma_m.$$

   Then by countable additivity,

$$\mu\left(\bigcup_{n=1}^{\infty} \Gamma_n\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) \leq \sum_{n=1}^{\infty} \mu(B_n) \leq \sum_{n=1}^{\infty} \mu(\Gamma_n).$$

In the last step we used $B_n \subseteq \Gamma_n$ and part 1.

3. Follows directly from part 2.

4. Like in part 2, take $A_n = B_n \backslash B_{n-1}$. Then

$$\mu(B_n) = \sum_{m=1}^{n} \mu(A_m) \nearrow \sum_{m=1}^{\infty} \mu(A_m) = \mu(B).$$

5. By the previous part, $B_1 \backslash B_n \nearrow B_1 \backslash B$ giving $\mu(B_1 \backslash B_n) \nearrow \mu(B_1 \backslash B)$.

   Now use $\mu(B \backslash A) = \mu(B) - \mu(A)$, which holds because $\mu(B) < \infty$.

$\square$

Note item 5 is false without the assumption that $\mu(B_1) < \infty$. For example, consider the measure on $\mathbb{Z}$ with $\mu(\Gamma) = |\Gamma|$, and take $B_n = \{i : i \geq n\}$.

Note from item 3, the existence of Lebesgue measure implies $\mathbb{R}$, or any interval of $\mathbb{R}$, is uncountable, since all countable subsets have measure 0 and any interval does not.

Definition 5.6: We say $\mathcal{C}$ is a Π-system if $\mathcal{C}$ is closed under intersection, i.e. if $A \in \mathcal{C}$ and $B \in \mathcal{C}$ then $A \cap B \in \mathcal{C}$.

Lecture 6
Mon. 2/14/2011

§1 More on σ-algebras

We give another characterization of $\sigma(\mathcal{C})$, the smallest σ-algebra containing $\mathcal{C}$.

Definition 6.1: We say that $\mathcal{H}$ is a Λ-system if

1. $E \in \mathcal{H}$.

2. If $A, B \in \mathcal{H}$ and $A \cap B = \phi$ then $A \cup B \in \mathcal{H}$.

3. If $A, B \in \mathcal{H}$ and $A \subseteq B$ then $B \backslash A \in \mathcal{H}$.

4. If $\{A_n : n \geq 1\} \subseteq \mathcal{H}$ and $A_n \nearrow A$ then $A \in \mathcal{H}$.

Theorem 6.2: Suppose $\mathcal{C}$ is a Π-system with $\mathcal{C} \subseteq \mathcal{P}(E)$. Let μ, ν on $\sigma(\mathcal{C})$ be such that $\mu(E) = \nu(E) < \infty$ and $\mu(A) = \nu(A)$ for all $A \in \mathcal{C}$. Then $\mu(A) = \nu(A)$ for all $A \in \sigma(C)$.

If two measures agree on the whole set E and a Π-system, then they agree on the smallest σ-algebra generated by the Π-system. (cf. Two continuous functions equal on a dense set are equal on the whole set.)

Proof. The set of subsets $\mathcal{H}'$ on which μ and ν agree satisfy conditions 1 (by assumption) and 2 (by additivity). It satisfies condition 3 because $\mu(B \backslash A) = \mu(B) - \mu(A)$ (measures are finite). It satisfies condition 4 by Proposition 5.5(4). Hence $\mathcal{H}'$ is a Λ-system. (This is the motivation for the definition of a Λ-system.)

It suffices to show that

$$\sigma(\mathcal{C}) = \bigcap \{\mathcal{H} : \mathcal{H} \text{ is } \Lambda\text{-system containing } \mathcal{C}\} =: \mathcal{H}_0.$$

(In other words $\sigma(C)$ is the smallest Γ-system containing $\mathcal{C}$.)

We first show that $\mathcal{H}_0$ is a σ-algebra.

Lemma 6.3: $\mathcal{B}$ is a σ-algebra iff $\mathcal{B}$ is both a Π and Λ-system.

Proof. The forward direction is clear. For the reverse direction, take $B = E$ in condition 3 to see $\mathcal{B}$ is closed under complementation. If $A, B \in \mathcal{B}$ then $A \cup B \in \mathcal{B}$, since we can write $A \cup B$ as a union of disjoint sets in $\mathcal{B}$ and use condition 2 as follows:

$$A \cup B = A \cup (B \backslash (A \cap B)).$$

Thus (by induction) $\mathcal{B}$ is closed under finite union. Now consider $\{A_n : n \geq 1\} \subseteq \mathcal{B}$. Then $\bigcup_{m=1}^{\infty} A_m \nearrow \bigcup_{m=1}^{\infty} A_m$ so by condition 4, $\bigcup_{m=1}^{\infty} A_m \subseteq \mathcal{B}$, and $\mathcal{B}$ is closed under countable union. $\square$

Now $\mathcal{H}_0$ is a Λ-system because it is the intersection of a family of Λ-systems. Now

$$\mathcal{H}_1 = \{\Gamma \subseteq E : \Gamma \cap A \subseteq \mathcal{H}_0 \text{ for every } A \in \mathcal{C}\}$$

is a Λ-system (check it!). Since $\mathcal{C}$ is a Π-system, $\mathcal{C} \subseteq \mathcal{H}_1$ and hence $\mathcal{H}_1 \supseteq \mathcal{H}_0$ ($\mathcal{H}_0$ being the smallest Λ-system containing $\mathcal{C}$). This gives

$$\Gamma \cap A \in \mathcal{H}_0 \text{ for every } \Gamma \in \mathcal{H}_0, \Delta \in \mathcal{C}. \tag{1}$$

Let

$$\mathcal{H}_2 = \{\Gamma \subseteq E : \Gamma \cap A \in \mathcal{H}_0 \text{ for every } A \in \mathcal{H}_0\}.$$

Then $\mathcal{H}_2$ is a Λ-system; it contains $\mathcal{C}$ by (1). Hence $\mathcal{H}_0 \subseteq \mathcal{H}_2$, and $\mathcal{H}$ is a Π-system. $\square$

Given $(E, \mathcal{B}, \mu)$, can we extend the measure to an even larger σ-algebra? Yes.

Definition 6.4: Define the completion of B with respect to μ as

$$\bar{\mathcal{B}}^{\mu} = \{\Gamma \subseteq E : \text{there exist } A, B \in \mathcal{B}, A \subseteq \Gamma \subseteq B, \mu(B \backslash A) = 0\}.$$

We can define a measure $\bar{\mu}$ on $\bar{\mathcal{B}}^{\mu}$ by

$$\bar{\mu}(\Gamma) = \mu(A).$$

(This is well-defined because if $A_i \subseteq \Gamma \subseteq B_i$ and $\mu(B_i \backslash A_i) = 0$ for $i = 1, 2$ then $\mu(A_1) \leq \mu(B_2) = \mu(A_2) \leq \mu(B_1) = \mu(A_1)$ so $\mu(A_1) = \mu(A_2)$.) Then $(E, \bar{\mathcal{B}}^{\mu}, \bar{\mu})$ is called the **completion** of $(E, \mathcal{B}, \mu)$.

This is again a σ-algebra: Indeed $A \subseteq \Gamma \subseteq B$ implies $B^c \subseteq \Gamma^c \subseteq A^c$ with $\mu(A^c \backslash B^c) = \mu(B \backslash A)$. Similarly it's closed under countable union.

Definition 6.5: Let $\mathcal{G}(E)$ denote the open sets of the topological space E, and let $\mathcal{B} = \sigma(\mathcal{G}(E))$ be the Borel algebra with measure μ. $\Gamma \subseteq E$ is **μ-regular** if for every $\varepsilon > 0$ there exists $F \in \mathcal{F}(E)$ such that $G \in \mathcal{G}(E)$, $F \subseteq \Gamma \subseteq G$ and $\mu(G \backslash F) < \varepsilon$.

Restrict choice of bread: Upper slice is open and bottom slice is closed. But we're more lenient about the middle: it doesn't have to be 0, just less than ε.

Proposition 6.6: A regular set is in the completion.

Proof. Take $G_n \supseteq \Gamma \supseteq F$ with the property that $\mu(G_n \backslash F_n) \le \frac{1}{n}$. Without loss of generality we may assume that the G_n are decreasing. (Replace G_n with $\bigcap_{m=1}^n G_m$.) Similarly we may assume that F_n are increasing. Let

$$D = \bigcap_{n=1}^{\infty} G_n, \quad C = \bigcup_{n=1}^{\infty} F_n$$

D is not necessarily open and C is not necessarily closed but both are Borel sets. Hence they are in $\mathcal{B}$ (as countable intersections/complements of elements in $\mathcal{B}$ are in $\mathcal{B}$, and open sets are in $\mathcal{B}$). $\square$

Given a topology E, let $G_\delta(E)$ be the set of countable intersections of open sets, and let $F_\sigma(E)$ be the set of countable unions of closed sets. If E is a metric space, the open sets are in $F_\sigma(E)$. closed under ctable unions Clsoed sets are in $G_\delta(E)$. countable intersections $F_{\sigma\delta}(E)$=take countable unions of elements in $F_\sigma(E)$. Ad infinitum. Beyond countably infinitely many times, get all Borel sets.

Lecture 7
Wed. 2/16/2011

§1 Constructing measures

Let (E, ρ) be a metric space, and $\mathcal{R} \subseteq \mathcal{P}(E)$ be a family of compact subsets. Let V be a map $V : \mathcal{R} \to [0, \infty)$. (Keep in mind the model that $E = \mathbb{R}^N$, $\mathcal{R}$ is the set of closed rectangles, and V the volume of the rectangles.) Suppose the following hold.

1. $\mathcal{R}$ is a Π-system, i.e. if $I, J \in \mathcal{R}$ then $I \cap J \in \mathcal{R}$.

2. $\phi \in \mathcal{R}$ and $V(\phi) = 0$.

3. If $I \subseteq J$ then $V(I) \le V(J)$.

4. Suppose $\{I_1, \ldots, I_n\} \in \mathcal{R}$ and $J \in \mathcal{R}$.

   (a) If $J \subseteq \bigcup_{m=1}^n I_m$ then $V(J) \le \sum_{m=1}^n V(I_m)$.
   (b) If $J \supseteq \bigcup_{m=1}^n I_m$ and the I_m's are non-overlapping then $V(J) \ge \sum_{m=1}^n V(I_m)$.

5. For all $I \in \mathcal{R}$ and all $\varepsilon > 0$, there exist $I, I' \in \mathcal{R}$ such that $I'' \subseteq I^\circ$, $I \in I'^\circ$ and $V(I') \le V(I'') + \varepsilon$.

6. For all open $G \in \mathcal{G}(E)$ there exists a sequence $\{I_n : n \ge 1\} \subseteq \mathcal{R}$ nonoverlapping sets such that $G = \bigcup_{m=1}^\infty I_n$.

Proposition 7.1: These properties hold for $E = \mathbb{R}^N$, $\mathcal{R}$ is the set of closed rectangles, and V the volume of the rectangles.

Proof. Item 4 holds by Lemma 1.2. Item 5 holds since we can enlarge or shrink $\mathcal{R}$ by a tiny bit. For item 6, consider a checkerboard of cubes of side length 2^{-n}:

$$\{k2^{-n} + [0, 2^{-n}]^N : k \in \mathbb{Z}^N\}$$

Take $m < n$ and $Q \in \mathcal{C}_m$, $Q' \in \mathcal{C}_n$. Either

1. The interiors of Q, Q' intersect, and $Q' \subseteq Q$, or

2. The interiors of Q, Q' do not intersect.

I.e. "either Q' is a descendant of Q or they are unrelated."

We use a greedy algorithm to stuff cubes in G. We show that in fact item 6 can be done with cubes of arbitrarily small length.

By splitting G into bounded parts we may assume G is bounded. Let $\delta > 0$ be given.

Let $G_0 = G$ and define the $G_k, \mathcal{A}_k$ inductively as follows. Let n_k be the smallest n such that there exist $Q \in \mathcal{C}_n$ for $Q \subseteq G_k$ and $2^{-n} < \delta$.

Let $\mathcal{A}_k = \{Q \in \mathcal{C}_{n_k} : Q \subseteq G\}$ and

$$G_k = G_{k-1} \backslash \bigcup_{Q \in \mathcal{C}_{n_k}} Q.$$

Let

$$\mathcal{A} = \bigcup_{n=1}^\infty \mathcal{A}_n.$$

It is clear that $\mathcal{A} \subseteq G$. Take a point $x \in G$. Take n sufficiently large; there's a cube from $\mathcal{C}_n$ such that $x \in \mathcal{C}_n$. Either that cube was chosen or one of its ancestors (which contains that cube) was chosen. □

Our goal is to prove the following.

Theorem 7.2: Given conditions (1)–(6), there exists a Borel measure such that $\mu(I) = V(I)$ for all $I \in \mathcal{R}$.

Proof. We will proceed in the following steps.

1. Define $\tilde{\mu}$ for all sets $\Gamma \subseteq E$ by
$$\tilde{\mu}(\Gamma) = \inf\left\{\sum_{n=1}^{\infty} V(I_n) : I_n \in \mathcal{R} \text{ and } \Gamma \subseteq \bigcup_{m=1}^{\infty} I_m\right\}.$$
Then $\tilde{\mu}$ is subadditive (Lemma 7.3).

2. $\tilde{\mu}$ of a countable family of nonoverlapping sets J_l of $\mathcal{R}$ is just the sum of the volumes $V(J_l)$. (A generalization of the fact that $\tilde{\mu}$ agrees with V.) (Lemma 7.4)

3. Give an alternate characterization of $\tilde{\mu}$ (Proposition 8.1):
$$\tilde{\mu}(\Gamma) = \inf\{\tilde{\mu}(G) : G \in \mathcal{G}(E) \text{ and } \Gamma \subseteq G\}.$$

4. Let $\mathcal{L}$ be the collection of $\Gamma \subseteq E$ such that for every $\varepsilon > 0$ there exists $G \supseteq \Gamma$ with $\tilde{\mu}(G \backslash \Gamma) < \varepsilon$. Then $\mathcal{L}$ is a σ-algebra (Theorem 8.2).

5. $\mu = \tilde{\mu}|\mathcal{L}$ is a measure on $\mathcal{L}$ (Thereom 8.4).

□

Lemma 7.3: $\tilde{\mu}$ is sub-additive, i.e.
$$\tilde{\mu}\left(\bigcup_{n=1}^{\infty} \Gamma_n\right) \leq \sum_{n=1}^{\infty} \tilde{\mu}(\Gamma_n).$$

Proof. This is another application of the $2^{-m}\varepsilon$ trick. Given $\varepsilon > 0$, for each m, choose $\{I_{m,n} : n \geq 1\} \subseteq \mathcal{R}$ such that
$$\Gamma_m \subseteq \bigcup_{n=1}^{\infty} I_{m,n} \text{ and } \sum_{n=1}^{\infty} V(I_{m,n}) \leq \tilde{\mu}(\Gamma_m) + 2^{-m}\varepsilon.$$
The collection $\{I_{m,n} : (m,n) \in \mathbb{N}^2\}$ covers $\bigcup_{m=1}^{\infty} \Gamma_m$.

Since all terms are nonnegative, we can write the sum as an iterated sum.
$$\sum_{(m,n)\in\mathbb{N}^2} V(I_{m,n}) = \sum_{m\in\mathbb{N}} \sum_{n\in\mathbb{N}} V(I_{m,n}) \leq \sum_{m=1}^{\infty} \left(\tilde{\mu}(\Gamma_m) + \varepsilon 2^{-m}\right) \leq \sum_{m=1}^{\infty} \tilde{\mu}(\Gamma_m) + \varepsilon.$$

The lemma follows upon combining this with
$$\sum_{(m,n)\in\mathbb{Z}} V(I_{m,n}) \geq \tilde{\mu}\left(\bigcup_{m=1}^{\infty} I_{m,n}\right).$$

□

Now we look for a σ-algebra $\mathcal{B}_\mu \subseteq \mathcal{B}_E$ such that $\tilde{\mu}$ on $\mathcal{B}_\mu$ is a measure. We need to check that $\tilde{\mu}(I) = V(I)$. In fact we check the following stronger statement.

Lemma 7.4: Let $\{J_1, \ldots\} \subseteq \mathcal{R}$ be a set of nonoverlapping rectangles. Then

$$\tilde{\mu}\left(\bigcup_{l=1}^{\infty} J_l\right) = \sum_{l=1}^{\infty} V(J_l).$$

Proof. First we check that equality holds when there are a finite number of J_l, $1 \leq l \leq L$. Note "$\leq$" holds because $\tilde{\mu}$ is the infimum over all covers and J_l form a cover for $\bigcup J_l$. We need to show "$\geq$." Let $\{I_m : m \geq 1\}$ with $\bigcup_{l=1}^{L} J_l \subseteq \bigcup_{m=1}^{\infty} I_m$. Choose I'_m containing I_m in its interior so that $V(I'_m) \leq V(I_m) + \varepsilon 2^{-m}$ (same trick!). Now $\bigcup_{l=1}^{L} J_l$ is compact as it is a finite union of compact sets. Since

$$\bigcup_{m=1}^{\infty} I_m^\circ \supseteq \bigcup_{l=1}^{L} J_l,$$

we can take a finite subcover of $\bigcup J_l$:

$$\bigcup_{m=1}^{n} I_m^\circ \supseteq \bigcup_{l=1}^{L} J_l.$$

Consider the cover $I_{l,m} = J_l \cap I'_m$ (so we can change order of summation). Note $I_{1,m}, \ldots, I_{L,m}$ are nonoverlapping and $I'_m \supseteq \bigcup_{l=1}^{L} I_{l,m}$. Then by condition 4,

$$\left(\sum_{m=1}^{n} V(I_m)\right) + \varepsilon \geq \sum_{m=1}^{n} V_m(I'_m) \overset{4}{\geq} \sum_{m=1}^{n} \sum_{l=1}^{L} V(I_{l,m}) \overset{4}{\geq} \sum_{l=1}^{L} V(J_l).$$

Now take the infimum to get $\tilde{\mu}\left(\bigcup J_l\right)$ on the left-hand side.

Now we extend to the infinite case. We want

$$\tilde{\mu}\left(\bigcup_{l=1}^{\infty} J_l\right) = \sum_{l=1}^{\infty} V(J_l)$$

Note "$\leq$" holds because $\tilde{\mu}$ is the infimum over all covers and J_l form a cover. We need to show "$\geq$." Use the finite case

$$\tilde{\mu}\left(\bigcup_{l=1}^{n} J_l\right) \geq \sum_{l=1}^{n} V(J_l)$$

and take $n \to \infty$ to get the equation above. $\square$

Lecture 8
Fri. 2/18/2011

§1 Constructing measures, continued

We continue to assume all 6 conditions given in the previous lecture.

By assumption 6, given $G \in \mathcal{G}(E)$ we can write $G = \bigcup_{m=1}^{\infty} I_m$. Lemma 7.4 says

$$\tilde{\mu}(G) = \sum_{m=1}^{\infty} V(I_m).$$

It immediately follows that if $G \cap G' = \phi$ then

$$\tilde{\mu}(G \cup G') = \tilde{\mu}(G) + \tilde{\mu}(G').$$

(Just put the two families of rectangles together.)

Proposition 8.1:

$$\tilde{\mu}(\Gamma) = \inf\{\tilde{\mu}(G) : G \in \mathcal{G}(E) \text{ and } \Gamma \subseteq G\}.$$

Proof. Now "$\le$" obviously holds.

We need to show that given $\bigcup_{m=1}^{\infty} I_m \supseteq \Gamma$ and $\varepsilon > 0$,

$$\tilde{\mu}(G) \le \sum_{m=1}^{n} V(I_m) + \varepsilon.$$

Take I'_m so that $I'^{\circ}_m \supseteq I_m$ and $V(I'_m) \le V(I_m) + 2^{-m}\varepsilon$. Now take

$$G = \bigcup_{m=1}^{\infty} I'^{\circ}_m.$$

□

We want a family of sets in $\mathcal{P}(E)$ that is a σ-algebra, such that the restriction of $\tilde{\mu}$ there is countably additive and hence a measure.

Let $\mathcal{L}$ be the collection of $\Gamma \subseteq E$ such that for every $\varepsilon > 0$ there exists $G \supseteq \Gamma$ with $\tilde{\mu}(G \backslash \Gamma) < \varepsilon$.

Theorem 8.2: $\mathcal{L}$ is a σ-algebra.

Proof.

1. Every open set is in $\mathcal{L}$:
$$\mathcal{G}(E) \subseteq \mathcal{L}.$$
2. $\mathcal{L}$ is closed under countable unions, by the $2^{-n}\varepsilon$ argument.

3. "Sets of measure 0" are in $\mathcal{L}$:

$$\tilde{\mu}(\Gamma) = 0 \implies \Gamma \in \mathcal{L}.$$

(Take an open set $U \supseteq \Gamma$ so that $\tilde{\mu}(\Gamma) \leq \varepsilon$.)

4. Compact sets are in $\mathcal{L}$. We use the following lemma:

Lemma 8.3: Suppose $K, K' \subset\subset E$ with $K \cap K'$ and $K \cap K' = \phi$. (The notation $\subset\subset$ means "compact subset of".) Then

$$\tilde{\mu}(K \cup K') = \tilde{\mu}(K) + \tilde{\mu}(K').$$

Proof. "$\leq$" holds by subadditivity.

We can find disjoint open subsets G, G' containing K, K'. Let H be an open set such that $H \supseteq K \cup K'$. Then

$$\begin{aligned}
\tilde{\mu}(H) &\geq \tilde{\mu}((G \cap H) \cup (G' \cap H)) \\
&\overset{7.4}{=} \tilde{\mu}(G \cap H) + \tilde{\mu}(G' \cap H) \\
&= \tilde{\mu}(K) + \tilde{\mu}(K')
\end{aligned}$$

Taking the infimum of the LHS gives $\tilde{\mu}(K \cup K')$ by Lemma 8.1. □

Now we show that if $K \subset\subset E$ then $K \in \mathcal{L}$. We need to show for $\varepsilon > 0$ there exists $G \supseteq K$ so $\tilde{\mu}(G \backslash K) < \varepsilon$. We claim $\tilde{\mu}(K) < \infty$. By assumption 6, we can write $K = \bigcup_{m=1}^{\infty} I_m$. For each I_m we can choose open I'_m so that $I'^{\circ}_m \supseteq I_m$ with $V(I'_m) \leq V(I_m)+1$. We can choose a finite cover: $K \subseteq \bigcup_{m=1}^{n} I'^{\circ}_m$. Then $\tilde{\mu}(K) \leq \sum_{m=1}^{n} V(I'^{\circ}_m)$.

We can choose $G \supseteq K$ so that $\tilde{\mu}(G) \leq \tilde{\mu}(K) + \varepsilon$, i.e. $\tilde{\mu}(G) - \tilde{\mu}(K) \leq \varepsilon$. Look at $G \backslash K$; it is open. (K is compact in a metric space, so closed.) Thus by assumption 6, we can write $G \backslash K = \bigcup_{m=1}^{n} I_m$. Now we show

$$\tilde{\mu}\left(K \cup \bigcup_{m=1}^{n} I_m\right) \leq \tilde{\mu}(G)$$

Now $K \cup \bigcup_{m=1}^{n} I_m$ is compact because it is a finite union of compact sets. They are disjoint so by the Lemma 7.4 the volume is

$$\tilde{\mu}(K) + \sum_{m=1}^{n} V(I_m).$$

Hence

$$\tilde{\mu}(G \backslash K) = \sum_{m=1}^{\infty} V(I_m) \leq \varepsilon.$$

So K is in $\mathcal{L}$.

5. Closed sets F are in $\mathcal{L}$. Indeed, write $E = \bigcup_{m=1}^{\infty} I_m$ (assumption 6). Given closed F, $F_n = G \cap \bigcup_{m=1}^{n} I_m$ is compact (because a closed set in a compact set is compact) and hence in $\mathcal{L}$ by item 3. Now $F = \bigcup_{n=1}^{\infty} F_n$, so by item 2 ($\mathcal{L}$ closed under countable union), $F \in \mathcal{L}$.

6. Countable unions of closed sets are in $\mathcal{L}$, i.e. $\mathcal{F}_\sigma(E) \subseteq \mathcal{L}$: Use item 5 and item 2 ($\mathcal{L}$ closed under countable union).

7. If $\Gamma \in \mathcal{L}$ then $\Gamma^c \in \mathcal{L}$. Given $\Gamma \in \mathcal{L}$, choose $G_n \supseteq \Gamma$ so that $\tilde{\mu}(G_n \backslash \Gamma) \le \frac{1}{n}$. Let $D = \bigcap_{n=1}^{\infty} G_n$. Then $D \in \mathcal{G}_\delta(E)$, $D \supseteq \Gamma$, and $\tilde{\mu}(D \backslash \Gamma) = 0$. Hence $D \backslash \Gamma \in \mathcal{L}$ by item 3. Now $\Gamma^c \backslash D^c = D \backslash \Gamma$. So
$$\Gamma^c = D^c \cup (\Gamma^c \backslash D^c) \in \mathcal{L}.$$
Note $D^c \in \mathcal{F}_\sigma(E)$ so $D^c \in \mathcal{L}$ by item 6. Hence $D \in \mathcal{F}_\sigma \subseteq \mathcal{L}$.

We've shown the three defining properties of a σ-algebra (Definition 5.1) in items 1 (E is open), item 2 (closed under countable union), and item 7 (closed under complements). □

Theorem 8.4: $\mu = \tilde{\mu} | \mathcal{L}$ is a measure on $\mathcal{L}$.

Proof. We need to show μ is countably additive. Since $\Gamma^c \in \mathcal{L}$, given $\Gamma \in \mathcal{L}$ and $\varepsilon > 0$ there exists $F \in \mathcal{F}(E), F \subseteq \Gamma$ so that $\tilde{\mu}(\Gamma \backslash F) < \varepsilon$ by the same trick.

Assume we have $\Gamma_n, n \ge 1$ mutually disjoint, relatively compact (i.e. having compact closure) sets in $\mathcal{L}$. We first prove countable additivity in this case. Given $K_n \subset\subset E$, for $K_n \subseteq \Gamma_n$. Thus
$$\mu\left(\bigcup_{m=1}^{\infty} \Gamma_m\right) \ge \mu\left(\bigcup_{m=1}^{n} \Gamma_m\right) \ge \mu\left(\bigcup_{m=1}^{n} K_m\right) = \sum_{m=1}^{n} \mu(K_m).$$
Now take K_n such that $\mu(\Gamma_n) \le \mu(K_n) + \varepsilon 2^{-n}$ and take $n \to \infty$.

The opposite inequality holds by subadditivity of $\tilde{\mu}$.

For the general case, choose $I_m \in \mathcal{R}$ so that $E = \bigcup_{n=1}^{\infty} I_n$, and let $A_1 = I_1$, $A_{n+1} = I_{n+1} \backslash \bigcup_{m=1}^{n} I_m$. Then the A_n's are mutually disjoint, relatively compact sets in $\mathcal{L}$, and so are $\Gamma_{m,n} = A_m \cap \Gamma_n$. Now use the previous part on the $\Gamma_{m,n}$. □

This finishes the proof of existence.

Lecture 9
Tue. 2/22/2011

§1 Uniqueness

Theorem 9.1: There exists a unique Borel measure μ such that $\mu(I) = V(I)$ for $I \in \mathcal{R}$. μ is regular, and $\mathcal{L} = \overline{\mathcal{B}}_E^{\mu}$ and $\overline{\mu} = \tilde{\mu} | \overline{\mathcal{B}}_E^{\mu}$.

Proof. We've proved existence; it remains to prove uniqueness. Suppose ν is another such measure. First suppose G is open; write $G = \bigcup_{m=1}^{\infty} I_m$ (nonoverlapping cover). Now

$$\nu(G) \le \sum_{m=1}^{\infty} \nu(I_m) = \sum_{m=1}^{\infty} \mu(I_m) = \mu(G).$$

For each m we can find a rectangle $I_m'' \subseteq I_m^\circ$ with $V(I_m) \le V(I_m'') + \varepsilon 2^{-m}$. The I_m'' are in the interiors of disjoint rectangles so are disjoint. By countable additivity,

$$\nu(G) \ge \sum_{m=1}^{\infty} \nu(I_m'') \ge \mu(G) - \varepsilon.$$

This shows that $\mu = \nu$ on open sets.

Next we show that $\mu = \nu$ on compact sets K. Choose open G such that $K \subset G$ and $\mu(G \backslash K)$ is finite. Then, since compact sets have finite measure under μ, and both G and $G \backslash K$ are open,

$$\nu(K) = \nu(G) - \nu(G \backslash K) = \mu(G) - \mu(G \backslash K) = \mu(K).$$

Next we show that $\mu = \nu$ on closed sets F. Write $E = \bigcup_{m=1}^{\infty} I_m$. Now $F_n = F \cap \bigcup_{m=1}^{n} I_m$ are compact, so $\nu(F_n) = \mu(F_n)$. Then $F_n \nearrow F$ so $\nu(F) = \mu(F)$.

Now take an arbitrary set $\gamma \in \mathcal{B}_E$. Given $\varepsilon > 0$, take F closed and G open so that $F \subseteq \Gamma \subseteq G$ and $\mu(G \backslash F) < \varepsilon$. Then

$$\mu(\Gamma) - \varepsilon \le \mu(F) = \nu(F) \le \nu(\Gamma) \le \nu(G) = \mu(G) \le \mu(\Gamma) + \varepsilon$$

so $\mu(\Gamma) = \nu(\Gamma)$.

□

§2 Invariance of measure

Corollary 9.2: Suppose $T : E \to E$ is a map such that $T^{-1}(I) \in \mathcal{R}$ (the inverse image of a rectangle is a rectangle) and $V(T^{-1}(I)) = V(I)$ for all rectangles. Then T is measurable and $T_*\mu = \mu$, where $T_*\mu(\Gamma) = \mu(T^{-1}\Gamma)$ (the pushforward measure).

Proof. Since $\nu = T_*\mu$ is a Borel measure assigning to a rectangle, $V(I)$, by uniqueness $\nu = \mu$. □

Let $E = \mathbb{R}^N$ and $\mathcal{R}$ be a rectangle in $\mathbb{R}^N$, and $V(I) = \mathrm{vol}(I)$. Denote this measure (the Lebesgue measure) by $\lambda_{\mathbb{R}^N}$. Fix $x \in \mathbb{R}^N$ and let $T_x y = x$. Since $V(T_x^{-1}(I)) = V(I)$, by Corollary 9.2 $(T_x)_*\lambda_{\mathbb{R}^N} = \lambda_{\mathbb{R}^N}$.

Theorem 9.3: If μ is a translation invariant Borel measure on $\mathbb{R}^N$ and $\mu([0,1]^N) = 1$, then $\mu = \lambda_{\mathbb{R}^N}$.

Proof. First, we claim that if I is a rectangle, $\mu(\partial I) = 0$. It suffices to prove a prove that a hyperplane has measure 0; it suffices to show that a rectangle of dimension $N-1$ with side length 1 has measure 0. If it has positive measure, then we can put a countable number of translates of them in $[0,1]^N$ and find $[0,1]^N$ has infinite measure, contradiction.

Now $\mu([0,2^{-n}]^N) = 2^{-nN}$ because we can tile $[0,1]^N$ with 2^N translates of it; they intersect only at their boundary which has measure 0. But we've showed that any open set can be covered by rectangles of this type, so μ agrees with $\lambda_{\mathbb{R}^N}$ on open sets, and by Theorem 9.1, $\mu = \lambda_{\mathbb{R}^N}$. □

Let $A = [a_{ij}]_{1\le i,j\le N}$. Define T_A by $(T_Ax)_i = \sum_j a_{ij}x_j$. Suppose A is nonsingular (nonzero determinant, T_A is invertible, onto). Note $T_A = T_{A^{-1}}^{-1}$.

Theorem 9.4:

$$\lambda_{\mathbb{R}^N}(T_A\Gamma) = |\det(A)|\lambda_{\mathbb{R}^N}(\Gamma).$$

Proof. Define $\mu(\Gamma) = \lambda_{\mathbb{R}^N}(T_A\Gamma)$. It is a Borel, translation invariant measure. Let

$$\alpha(A) = \lambda_{\mathbb{R}^N}(T_A([0,1]^n)).$$

Now $\frac{\mu(\Gamma)}{\mu([0,1]^n)}$ is a translation-invariant Borel measure assigning 1 to the unit cube, so

$$\lambda_{\mathbb{R}^n}(\Gamma) = \alpha(A)\lambda_R(\Gamma)$$

Now $\alpha(A\circ B) = \alpha(A)\alpha(B)$.

There are two types of matrices where $\alpha(A)$ can be easily computed.

1. If A is a diagonal matrix $A = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$ then $\alpha(A) = \lambda_1\cdots\lambda_n = |\det(A)|$ because the cube is "stretched".

2. If A is orthogonal then $\alpha(A) = 1$ because it takes the unit ball to itself.

$$\alpha(0)\lambda_{\mathbb{R}^N}(B(0,1)) = \lambda_{\mathbb{R}^N}(B(0,1)) \implies \alpha(0) = 1.$$

If A is symmetric, then we can write $A = O^TDO$ where O is orthogonal and D is diagonal. Hence $\alpha(A) = |\det(A)|$

For general A, use the polar decomposition $A = QH$ where Q is orthogonal and H is symmetric to get $\alpha(A) = |\det(H)| = |\det(A)|$.

If A is singular then $\mathbb{R}^N$ is mapped to a subspace of dimension $n-1$; an orthogonal transformation brings into $\mathbb{R}^{N-1}\times\{0\}$, which has measure 0. □

Lecture 10
Wed. 2/23/2011

§1 Some cool stuff

Lemma 10.1 (Vitalli)**:** Suppose $\Gamma \subseteq \mathbb{R}$ is Lebesgue measurable with positive measure. Then

$$\Gamma - \Gamma = \{x - y : x, y \in \Gamma\} \supseteq [-\delta, \delta]$$

for some $\delta > 0$.

Proof. Without loss of generatlity, we can assume Γ has finite Lebesgue measure. Then there exists $G \in \mathcal{G}(\mathbb{R})$ such that $\Gamma \subseteq G$ and $\lambda_{\mathbb{R}}(G \backslash \Gamma) \leq \frac{1}{3}\lambda_{\mathbb{R}}(\Gamma)$. Then

$$\lambda(\Gamma) \geq \frac{3}{4}\lambda(G).$$

Every open set of the real line can be written as the union of disjoint open intervals; write $G = \bigcup_{m=1}^{\infty} I_m$. Then

$$\sum_{m=1}^{\infty} \lambda(\Gamma \cap I_m^\circ) = \lambda(\Gamma) \geq \frac{3}{4}\lambda(G) = \frac{3}{4}\sum_{m=1}^{\infty} \lambda(I_m^\circ)$$

For at least one open interval,

$$\lambda(\Gamma \cap I_m^\circ) \geq \frac{3}{4}\lambda(I_m^\circ),$$

i.e. Γ looks large in some open interval. Let $A = \Gamma \cap I_m^\circ$ and suppose $A \cap (d + A) = \phi$. (Note A intersects $d + A$ iff $d \in A - A$.) Then

$$\frac{3}{2}\lambda(I_m^\circ) \leq 2\lambda(A) = \lambda(A \cup (d + A)).$$

Now $(a, b) \cup (a + d, b + d)$ is contained in an interval of length $b - a + |d|$ (either $(a, b + d)$ or $(a + d, b)$). Then

$$|d| \geq \frac{1}{2}\lambda(I_m^\circ).$$

So we can let $\delta = \frac{1}{2}\lambda(I_m^\circ)$. □

Theorem 10.2: Assuming the Axiom of Choice, given any Γ be any set of positive measure, there exists a nonmeasurable set contained in Γ.

Proof. Define an equivalence relation by

$$x \sim y \iff x - y \in \mathbb{Q}.$$

Using the Axiom of Choice, let A be a set with exactly one representative from each equivalence class.

Let Γ be any Lebesgue measurable set with positive measure. Consider the sets $\Gamma \cap (q+A)$. Suppose the intersection were Lebesgue measurable for every $q \in \mathbb{Q}$. But

$$\Gamma = \bigcup_{q \in \mathbb{Q}} (\Gamma \cap (q + A)) = \Gamma.$$

If Γ has positive measure, then there exists q such that $\Gamma \cap (q + A)$ is Lebesgue measurable of positive measure. By Vitalli's Lemma 10.1, $\Gamma \cap q + A$ contains an entire interval. But this is a contradiction because the difference of two elements in $q + A$ is rational only if it is zero. $\square$

§2 Probability measures

Theorem 10.3: Let $F : \mathbb{R} \to \mathbb{R}$ be a function. Then there exists a finite Borel measure on $\mathbb{R}$ such that

$$\mu_F((-\infty, x]) = F(x)$$

for all $x \in \mathbb{R}$ if and only if

1. F is increasing.
2. F is bounded.
3. $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$.
4. F is right continuous.

F is called a **distribution function** (for the measure μ_F).

Proof. Suppose μ is a finite Borel measure on $\mathbb{R}$. Define $F(x) = \mu((-\infty, x])$. Then

1. F is nondecreasing because $x \leq y$ implies $(-\infty, x] \subseteq (-\infty, y]$.
2. F is bounded because μ is finite.
3. $F(-\infty) = 0$ because the the measure of a decreasing sequence of sets $(-\infty, -n]$ is the limit of the measures.
4. F is right continuous for the same reason because

$$(-\infty, x] = \bigcap_{n=1}^{\infty} \left(-\infty, x + \frac{1}{n}\right].$$

Given F satisfying the conditions we can write $F = F_d + F_c$ where F_d is pure jump and F_c is nondecreasing. Thus we just need to consider F pure jump or nondecreasing.

For F_d pure jump, let D be the set of discontinuities. Then

$$F_d(x) = \sum_{y \in D \cap (-\infty, x]} (F_d(x) - F_d(x-)).$$

This motivates us to take

$$\mu_{F_d}(\Gamma) = \sum_{y \in \Gamma \cap D} (F_d(y) - F_d(y-)).$$

(Just add up the jumps in the set.)

For F_c continuous, let R be the set of closed intervals of the real line. Define V by

$$V([a,b]) = \mu_{F_c}([a,b]).$$

We check this satisfies hypotheses, so μ_{F_c} defines a Borel measure.

(Assumptions 4 and 5 come from continuity; for assumption 4 emulate the proof of Lemma 1.2.)

Now we show uniqueness. Suppose $\nu((-\infty, x]) = F(x)$. It is finite, and by taking a sequence of intervals increasing to (a,b),

$$\nu_F((a,b)) = \sum_{n \to \infty} \left\{ F\left(b - \frac{1}{n}\right) - F(\alpha) \right\} = F(b-) - F(a).$$

Every open set is the union of disjoint intervals; $\nu = \mu_F$ on open intervals so they agree on open sets. Then $\nu = \mu_F$. □

We talk about coin tossing. Suppose each coin is heads independently with probabity $p \in (0,1)$; let $q - 1 = a$. Let $(\eta_1, \dots, \eta_n) \in \{0,1\}^n$. The probability is $p^{\sum_{i=1}^n \eta_i} q^{n - \sum_{i=1}^n \eta_i}$.

We want to consider infinite coin tossing games, and describe events that depend on infinite number of tosses. We want to find a measure! To be continued.

Lecture 11
Fri. 2/25/2011

§1 Bernoulli measure

Let $\Omega = \{0,1\}^{\mathbb{N}}$, i.e. the set of binary sequences. Think of Ω as containing all possible outcomes of an infinite number of coin tosses. For $S \subseteq \mathbb{N}$, let

$$\Omega(S) = \{0,1\}^S;$$

think of this as containing all possible outcomes of games played during S. Define $\Pi_S : \Omega \to \Omega(S)$ as the restriction map.

Let

$$\mathcal{A}(S) = \{\Pi_S^{-1}(\Gamma) : \Gamma \subseteq \Omega(S)\},$$

the set of inverse images under the projection (collection of subsets of Ω). These are events that can be described entirely by outcomes of flips in S. Let $\eta \in \Omega(S)$. We assign the probability (probability of heads is p, tails is $q = 1 - p$, independent events)

$$P(\Pi_S^{-1}(\{\eta\})) = p^{\sum_{i\in F}\eta_i} q^{n-\sum_{i\in F}\eta_i}. \tag{2}$$

If S is infinite this is 0 (assuming $p \in (0,1)$). Hence we restrict our attention to finite subsets first. Let $A \in \mathcal{A}(F)$, $A = \prod_F^{-1}\Gamma$, $\Gamma \subseteq \Omega(F)$. Since

$$A = \bigsqcup_{\eta\in\Gamma} \Pi_F^{-1}(\{\eta\}),$$

we assign the following probability

$$P(A) = \sum_{\eta\in\Gamma} P(\Pi_S^{-1}(\{\eta\})) = \sum_{\eta\in\Gamma} p^{\sum_{i\in F}\eta_i} q^{n-\sum_{i\in F}\eta_i}.$$

Theorem 11.1: There exists a measure β_p such that $\beta_p(A) = P(A)$ for A in the form (2).

Proof. Suppose F_1, F_2 are finite subsets of $\mathbb{N}$, $\Gamma_1 \subseteq \Omega(F_1), \Gamma_2 \subseteq \Omega(F_2)$ and $\Pi_{F_1}^{-1}\Gamma_1 = \Pi_{F_2}^{-1}\Gamma_2$. (For example, $F_1 = \{1\}, F_2 = \{1,2\}$, $\Gamma_1 = \{1\}$, $\Gamma_2 = \{(1,0),(1,1)\}$; both $\Pi_{F_1}^{-1}\Gamma_1, \Pi_{F_2}^{-1}\Gamma_2$ equal the event that we get heads first toss.)

Without loss of generality, $F_1 \subseteq F_2$, and

$$P(\Pi_{F_1}^{-1}\Gamma_1) = \frac{|\Gamma_1|}{2^{|F_1|}} = \frac{2^{|F_2|-|F_1|}|\Gamma_1|}{2^{|F_2|}} = \frac{|\Gamma_2|}{2^{|F_2|}} = P(\Pi_{F_2}^{-1}\Gamma_2).$$

Thus P is well-defined on $\mathcal{A} = \bigcup_{F\subset\subset\mathbb{N}} \mathcal{A}(F)$.

Note $\mathcal{A}$ is closed under complementation and finite union (i.e. is an algebra). We can verify

$$\beta_p(A_1 + A_2) = \beta_p(A_1) + \beta_p(A_2) :$$

just take F so A_1, A_2 depend only on tosses in F.

To carry out our program (Theorem 7.2 and 9.1), we need to put a topology on Ω. We want $\omega_n \to \omega$ if $\omega_n(i) \to \omega(i)$ for al i, so we metrize Ω with

$$\rho(\omega, \omega') = \sum_{i\in\mathbb{N}} 2^{-1}|\omega(i) - \omega'(i)|.$$

Now (Ω, ρ) is compact: just show sequential compactness; given $\omega_n, n \geq 1$, by the box principle take a subsequence $\omega_{1,n}, n \geq 1$ so that $\omega_{1,n}(1)$ is constant. Take a subsequence $\omega_{2,n}$ so it's constant at time 2, and so on. Then look at the diagonal $\omega_{n,n}$. (Alternatively, note Ω is the product topology, and Ω is complact by Tychonoff's Theorem.)

Then for F finite and $\Gamma \subseteq \Omega(F)$, $\Pi_F^{-1}\Gamma$ is closed, and hence compact. But $\Pi_F^{-1}(\Omega(F)-\Gamma)$ is also closed so $\Pi_F^{-1}\Gamma$ is open.

Now we check the assumptions for our program ($V = P$): β_p is additive for disjoint sets so subadditive on sets. For V there exist $A'^{\circ} \subseteq A$ and $A^{\circ} \subseteq A''^{\circ}$ with close measure—just take $A = A' = A''$! Every open set can be written as a nonoverlapping (here, disjoint) union of countable number of A's: use the same greedy algorithm as in Proposition 7.1.

Then Theorem 9.1 applies. $\square$

$\beta_{\frac{1}{2}}$ models tossing with a fair coin. Let

$$\Phi(\omega) = \sum_{i\in\mathbb{N}} \frac{\omega(i)}{2^i} \in [0,1].$$

For $B \in \mathcal{B}_{[0,1]}$,

$$\begin{aligned} \beta_{\frac{1}{2}}(\{\omega : \Phi(\omega) \in B\}) &= \lambda_{\mathbb{R}}(B) \\ \beta_{\frac{1}{2}}(\Phi^{-1}(\Gamma) &= \lambda_{[0,1]}(\Gamma)). \end{aligned}$$

(Binary sequences with random coefficient.)

In other words,

$$\Phi_*\beta_{\frac{1}{2}} = \lambda_{[0,1]}.$$

Note that ω is not one-to-one, because $\omega(m) = 0$ and $\omega(i) = 1$ for $i \ge m+1$ gets mapped to the same thing with $\omega(m) = 1$ and $\omega(i) = 0$ for $i \ge m+1$ instead. So let

$$\hat{\Omega} = \{\omega : \omega(i) = 0 \text{ for infinitely many } i's.\} \cup \{\omega_\infty\}$$

where ω_∞ is the all 1's sequence.

See text for more.

Lecture 12
Mon. 2/28/2011

§1 Lebesgue Integration

Definition 12.1: Let (E, B, μ) be a measure space. f is **measurable** if

$$f^{-1} \in \mathcal{B} \text{ for all } f \in \mathcal{B}_R.$$

We want to define

$$\int f\, d\mu = \int f(x)\mu(dx).$$

The idea is to carve up the space so f nearly constant on each part. The simplest f are those taking a finite number of values; we can carve up the domain into finitely many pieces.

Definition 12.2: $f : E \to R$ is **simple** if $f(E)$ is finite. Then we define

$$\int f\, d\mu = \sum_{a \in f(E)} a\mu(\{f = a\}).$$

Careful with ∞... $\infty - \infty$ gives a problem. So we restrict to $f \geq 0$ for now.

The advantage of Riemann's Theory is that we do carving before looking at f, but we cab only integrate a restricted set of f's. Here we do carving afterwards, so properties are less obvious.

Lemma 12.3: Let $f : E \to [0, \infty)$ be measurable simple and $f = \sum_{l=1}^{\infty} 1_{\Delta_l}$. Then

$$\int f\, d\mu = \sum_{l=1}^{n} \beta_l \mu(\Delta_l).$$

Proof. Chop the domain into little pieces and expand the sums in terms of those little pieces. On each piece the assertion will be trivial.

1. Let $f(E) = \{a_1, \ldots, a_m\}$, $\Lambda_k = \{x : f(x) = a_k\}$. Refining the sum on the RHS gives

$$\sum_{l=1}^{n} \beta_l \mu(\Delta_l) = \sum_{k=1}^{m} \sum_{l=1}^{n} \beta_l \mu(\Delta_l \cap \Gamma_k).$$

It suffices to focus on one Γ_k, so we may assume $f = a1_\Gamma$.

2. The n sets Δ_n divide the points of space into 2^n categories. Namely, for $\eta \in \{0, 1\}$, we define $\Delta_\eta = \bigcap_{l=1}^{n} \Delta_l^{\eta_l}$ where $\Delta_l^1 = \Delta_l$ and $\Delta_l^0 = \Delta_l^c$. Let $\beta_\eta = \sum_{l=1}^{n} \eta_l \beta_l$, i.e. the sum of the β_l's that contribute to the value on Δ_η. Then refining the sum over the Δ_η,

$$\sum_{\eta \in \{0,1\}^n} \beta_\eta 1_{\Delta_\eta} = \sum_{l=1}^{n} \beta_l \mu(\Delta_l) = \alpha 1_\Gamma.$$

Since the Δ_η are disjoint, $\beta_\eta = \alpha 1_\Gamma$.

3. Then

$$\sum_{l=1}^{n} \beta_l \mu(\Delta_l) = \sum_{l=1}^{n} \beta_l \sum_{\{\eta : \eta_l = 1\}} \mu(\Delta_n) = \sum_{\eta \in \{0,1\}^n} = \alpha \mu(\Gamma) = \int_\Gamma f.$$

□

Proposition 12.4: For any α, β,

$$\int (\alpha f + \beta g)\, d\mu = \alpha \int f\, d\mu + \beta \int g\, d\mu.$$

Proof. Just write f and g in the form $\sum_{k=1}^{m} a_k 1_{A_k}$ and $\sum_{k=1}^{n} b_k 1_{B_k}$ and use the lemma. □

Lecture 13
Wed. 3/2/2011

For a simpler proof of Monday's stuff see updated text.

§1 Lebesgue Integration

Let $\bar{\mathbb{R}} = [-\infty, \infty]$ (a two-point compactification of $\mathbb{R}$). By convention,

$$0 \cdot \infty = 0, \quad \pm\infty \mp \infty = \text{undef.}$$

Let $\widehat{\bar{\mathbb{R}}^2} = \bar{\mathbb{R}}^2 \backslash \{(-\infty, \infty), (\infty, -\infty)\}$. We now let functions take values in $\bar{\mathbb{R}}$.

Corollary 13.1: For simple functions $f \le g$,

$$\int f \, d\mu \le \int g \, d\mu.$$

If $\int f \, d\mu < \infty$ then

$$\int g - f \, d\mu = \int g \, d\mu - \int f \, d\mu.$$

Proof. Write $g = f + (g - f)$ for the first assertion. □

Now we complete the definition of Lebesgue integral for non-simple functions. Let $f \ge 0$ be a measurable function. Let $\{\varphi_n : n \ge 1\}$ be a nondecreasing sequence of nonnegative simple functions such that $\varphi_n \nearrow f$ and define

$$\int f = \lim_{n \to \infty} \varphi_n.$$

For instance, take

$$\varphi_n(x) = \begin{cases} m2^{-m}, & \text{if } m2^{-n} \le f(x) < (m+1)2^{-n} \text{ for } m \le 4^n. \\ 2^n, & \text{if } f(x) \ge 2^n. \end{cases}$$

The following gives that the integral is well-defined.

Lemma 13.2: Suppose $\{\varphi_n\}$ is a sequence of nonnegative simple functions, and $\lim_{n \to \infty} \varphi_n \ge \psi$. Then

$$\lim_{n \to \infty} \int \varphi_n \, d\mu \ge \int \psi_n \, d\mu.$$

If we had $\varphi_n \nearrow f$ and $\psi_n \nearrow f$, we get $\lim_{n \to \infty} \varphi_n \ge \psi_m$ so taking the limit we get $\lim_{n \to \infty} \varphi_n \ge \lim_{n \to \infty} \psi_n$. Similarly with φ_n and ψ_n reversed. The lemma then shows that the integral is well-defined.

Proof. Consider 3 cases.

1. $\mu(\psi = \infty) > 0$: Now for $M < \infty$,

$$\lim_{n\to\infty} \mu(\{x : \varphi_n(x) > M\}) = \mu\left(\bigcup_n \{x : \varphi_n(x) > M\}\right) \geq \mu(\{x : \psi(x) = \infty\}).$$

   (increasing sequence of sets! Proposition 5.5(4)(*)). Thus by (13.1),

$$\varphi_n \geq M1_{\{\varphi_n > M\}}$$
$$\lim_{n\to\infty} \int \varphi_n \geq M\mu(\{\varphi = \infty\}),$$

   showing $\int \varphi_n \, d\mu \to \infty$.

2. $\mu(\psi > 0) = \infty$. Since ψ is simple, actually $\mu(\psi > \varepsilon) = \infty$ for some $\varepsilon > 0$. Then $\lim_{n\to\infty} \int \varphi_n = \infty = \int \psi$.

3. $\mu(\psi = \infty) = 0$ and $\mu(\psi > 0) < \infty$. Let $\hat{E} = \{0 < \psi < \infty\}$; this has finite measure. Now

$$\int_{\hat{E}} \psi = \int \psi.$$

   (The set where $\psi = \infty$, $\psi = 0$ make no contribution, the first because μ is 0 and the latter because $\psi = 0$.) Now

$$\int \varphi_n \geq \int_{\hat{E}} \varphi_n.$$

   (Integrate over less.) We need to show

$$\lim_{n\to\infty} \varphi_n \geq \int_{\hat{E}} \psi.$$

   Thus we can assume that $0 < \psi < \infty$. Since ψ is a simple function there exist $\varepsilon > 0$, $M < \infty$ so that $\varepsilon \leq \psi \leq M$. Consider $0 < \delta < \varepsilon$. Then

$$\begin{aligned}
\int \varphi_n &\geq \int_{\{\varphi_n \geq \psi - \delta\}} \varphi_n \\
&\geq \int_{\{\varphi_n \geq \psi - \delta\}} (\psi - \delta) \\
&= \left(\int_{\{\varphi_n \geq \psi - \delta\}} \psi\right) - \delta\mu(\{\varphi_n \geq \psi - \delta\}) \\
&\geq \int \psi - \int_{\{\varphi_n < \psi - \delta\}} -\delta\mu(E) \\
&\geq \int \psi - M\mu(\varphi_n < \psi - \delta) - \delta\mu(E).
\end{aligned}$$

The sets $\{\varphi_n \leq \psi - \varphi\}$ have finite measure and decrease to the empty set, so their measures decrease to 0 (*). Taking $n \to \infty, \delta \to 0$,

$$\lim_{n\to\infty} \varphi_n \geq \int \psi.$$

□

This lemma uses the full power of countable additivity. We used it crucially twice (*).

Proposition 13.3: For all nonnegative measurable functions,

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

If $f \leq g$ then $\int f \leq \int g$.

Theorem 13.4 (Markov Inequality)**:** Let $f \geq 0$. Then

$$\mu(f \geq R) \leq \frac{1}{R} \int_{\{f \geq R\}} f \leq \frac{1}{R} \int f.$$

Proof. Integrate

$$\lambda 1_{f\geq\lambda} \leq 1_{f\geq\lambda} f \leq f$$

and divide by R. □

Corollary 13.5: If $\int f = 0$ then $\mu(f > 0) = 0$.

Proof. By Proposition 5.5 and Markov's inequality,

$$\mu(f > 0) = \lim_{n\to\infty} \mu\left(f \geq \frac{1}{n}\right) = \lim_{n\to\infty} 0 = 0.$$

□

One more step in defining Lebesgue integral: Get away from dependence on nonnegativity. There is a canonical way of writing f as a difference of a nonnegative and nonpositive function, $f = f^+ - f^-$. Define

$$\int f = \int f^+ - \int f^-$$

when one of the two integrals on the RHS is finite. We have to check the integral is linear, $\int (f+g) = \int f + \int g$. (Assume $\int f, \int g$ are defined and they are not infinite with opposite sign, i.e. $(\int f^\pm) \vee (\int g^\pm) < \infty$.) Take $E^+ = \{f + g \geq 0\}$ and $E^- = \{f + g < 0\}$. Then

$$\begin{aligned}\int_{E^+} (f+g) &= \int_{E^+} [(f^+ + g^+) - (f^- + g^-)]\, d\mu \\ &= \int_{E^+} (f^+ + g^+) - \int_{E^+} (f^- + g^-) \\ &= \int_{E'} f^+ + \int_{E^+} g^+ - \int_{E^+} f^- - \int_{E^+} g^-.\end{aligned}$$

and similarly for E^-. Add everything up!

A recap:

1. We first define integrals for simple functions.
2. We define integrals for increasing limit of simple functions.
3. We define integrals for differences of two nonnegative functions.

Lecture 14
Fri. 3/4/2011

§1 Convergence Theorems

Definition 14.1: Let $f : E \to \bar{\mathbb{R}}$. Define the L^1 **norm** of f to be

$$||f||_{L'} = \int |f| \, d\mu.$$

If this is finite we say that f is μ-integrable.

The set of μ-integrable functions is denoted by $L^1(\mu; \mathbb{R})$.

Since $\mu(f = \infty) = 0$ is required, we assume without further comment that f is a function $E \to \mathbb{R}$.

Proposition 14.2: $L^1(\mu; \mathbb{R})$ is a vector space. If $f, g \in L^1$ then

$$||\alpha f + \beta g||_{L^1} \le |\alpha| \cdot ||f||_{L^1} + |\beta| \cdot ||g||_{L^1}.$$

Proof. Take integrals of $|af + bg| \le |\alpha||f| + |\beta||g|$. $\square$

Note $d(f, g) := ||f - g||_{L'}$ satisfies the following.

1. (Nonnegativity) $||f - g||_{L^1} \ge 0$.
2. (Triangle inequality) $||f - g||_{L^1} \le ||f - h||_{L^1} + ||h - g||_{L^1}$.

Note that d is not a metric, because $||f - g||_{L^1} = 0$ does not imply that $f = g$; we can only say $\mu(f \ne g) = 0$. Hence we identify f with g, and write $f \overset{\mu}{\sim} g$, if $\mu(f \ne g) = 0$. ($f = g$ almost everywhere.) The equivalence classes are

$$[f]^{\overset{\mu}{\sim}} = \{g : g \overset{\mu}{\sim} f\}.$$

We view $L^1(\mu; \mathbb{R})$ as the quotient space under this equivalence relation. Then $||[f]^{\overset{\mu}{\sim}}|| = ||f||_{L^1}$ is a norm.

Proposition 14.3: 1. (Pointwise limits of measurable functions are measurable.) Suppose $\{f_n : n \geq 1\}$ is a sequence of functions $(E, \mathcal{B}) \to (F, \mathcal{B}_F)$ such that for all $x \in E$, $\lim_{n\to\infty} f_n(x) = f(x)$. Then f is measurable.

2. Assume F admits a complete metric d. Let

$$\Delta = \{x \in E : \lim_{n\to\infty} f_n(x) \text{ exists}\}.$$

Let $y_0 \in F$. Define

$$f(x) = \begin{cases} \lim_{n\to\infty} f_n(x), & x \in \Delta \\ y_0, & x \notin \Delta. \end{cases}$$

Then f is measurable.

3. Let $F = \bar{\mathbb{R}}$. Then $\sup f_n, \inf f_n, \limsup f_n, \liminf f_n$ are measurable.

Proof. 1. Let $\Gamma \in \mathcal{B}_F$. Then

$$\{\Gamma : f^{-1}(\Gamma) \in \mathcal{B}\}$$

is a σ-algebra, because f^{-1} preserves unions and complements.

Since $\mathcal{B}_F$ is generated by open sets, we only need to verify that if G is open, then $f^{-1}(G) \in \mathcal{B}$. We show that

$$\{x : \lim_{n\to\infty} f_n(x) \in G\} \in \mathcal{B}.$$

We rewrite the set as a countable union/intersection of measurable sets, from which the conclusion is clear.

$$\bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \{x : f_n(x) \in G\}.$$

2. First we show Δ is measurable. Δ is the set of x so that $f_n(x)$ forms a Cauchy sequence. Then the set can be written as

$$\Delta = \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \left\{ d(f_n, f_m) < \frac{1}{k} \right\}$$

The terms are measurable because $\{d(x, y) < \frac{1}{k}\}$ is open.

Now suppose $\Gamma \in \mathcal{B}_F$. Then

$$f^{-1}(\Gamma) = \begin{cases} \{\lim_{n\to\infty} f_n(x) \in \Gamma\}, & y_0 \notin \Gamma \\ \Delta^c \cup \{\lim_{n\to\infty} f_n(x) \in \Gamma\}, & y_0 \in \Gamma. \end{cases}$$

3. Write as a limit of measurable functions.

$$\sup f_n = \lim_{n\to\infty}(f_1 \vee \cdots \vee f_n).$$
$$\limsup f_n = \lim_{n\to\infty}\sup_{n\geq m} f_n.$$

Similarly for inf and liminf.

□

Theorem 14.4 (Monotone Convergence Theorem)**:** Let $\{f_n : n \geq 0\}$ be a sequence of nonnegative measurable functions $(E, \mathcal{B}, \mu)$ such that $f_n \nearrow f$. Then

$$\lim_{n\to\infty}\int f_n\,d\mu = \int f\,d\mu.$$

I.e. we can pass the integral inside the limit.

In particular, if $f \in L^1$ then $||f_n - f||_{L^1} \to 0$ as $n \to \infty$.

Proof. The "$\leq$" is clear.

Go back to the definition. Given $\psi_n \geq 0$ simple so that $\psi_n \nearrow f$, then $\int f\,d\mu = \lim_{n\to\infty}\int \psi_n\,d\mu$. We want to link the ψ_n to the f_n.

$$\begin{array}{ccccc}
f_1 & f_2 & \to & f_n & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\varphi_{\mathbf{1},\mathbf{n}} & \varphi_{\mathbf{1},\mathbf{n}} & \cdots & \varphi_{\mathbf{n},\mathbf{n}} & \cdots \\
\uparrow & \uparrow & \cdots & \uparrow & \cdots \\
\varphi_{\mathbf{1},\mathbf{2}} & \varphi_{\mathbf{2},\mathbf{2}} & \cdots & \varphi_{n,2} & \cdots \\
\varphi_{\mathbf{1},\mathbf{1}} & \varphi_{2,1} & \cdots & \varphi_{n,1} & \cdots
\end{array}$$

We link the f_n to simple functions and use a "diagonal convergence" argument. Choose simple nonnegative $\varphi_{m,n}$ so that $\varphi_{m,n} \nearrow f_m$ as $n \to \infty$. Let

$$\psi_n = \max\{\varphi_{m,n} : 1 \leq m \leq n\}.$$

Now $\psi_n \leq \psi_{n+1}$ and $\varphi_{m,n} \leq \psi_n \leq f_n$. Let $n \to \infty$. Taking limits $f_m \leq \lim \psi_n \leq f$. Now let $m \to \infty$ to get $\lim \psi_n = f$. Hence, taking the integral of both sides,

$$\int f\,d\mu = \lim_{n\to\infty}\int \psi_n\,d\mu \leq \lim_{n\to\infty}\int f_n\,d\mu.$$

Now

$$||f - f_n||_{L^1} = \int f\,d\mu - \int f_n\,d\mu \to 0.$$

□

Theorem 14.5 (Fatou's Lemma)**:** Let $\{f_n : n \geq 1\}$ be measurable on $(E, \mathcal{B}, \mu)$.

1. If $f_n \geq 0$ then
$$\int \underline{\lim}\, f_n \, d\mu \leq \underline{\lim} \int f_n \, d\mu.$$

2. If there exists $g \in L^1$ such that $f_n \leq g$ then
$$\int \overline{\lim}\, f_n \, d\mu \geq \overline{\lim} \int f_n \, d\mu.$$

Proof. 1. Write
$$\underline{\lim}\, f_n = \lim_{m\to\infty} \underbrace{\inf_{n\geq m} f_n}_{g_m}.$$

Now $g_m \nearrow \underline{\lim}\, f_n$. By Monotone Convergence Theorem,
$$\underline{\lim} \int f_n \, d\mu \geq \lim_{m\to\infty} \int g_n \, d\mu = \int \underline{\lim}\, f_n \, d\mu.$$

2. Use the first part on $h_n = g - f_n$.

□

Lecture 15
Mon. 3/7/11

§1 Lebesgue dominated convergence

Suppose we have a measure space $(E, \mathcal{B}, \mu)$. We say that a property holds almost everywhere with to μ if it holds except on a set of measure 0. For example, $f = g$ (a.e., μ) means that $\mu(f \neq g) = 0$.

Theorem 15.1 (Lebesgue Dominated Convergence Theorem)**:** Suppose $\{f_n : n \geq 1\} \cup \{f\}$ is measurable, and there exists $g \in L^1(\mu)$ such that

1. for all n, $|f_n| \leq g$ almost everywhere w.r.t. μ, and

2. $f_n \to f$ almost everywhere,

Then $\left|\int f_n - \int f\right| \leq \int |f_n - f| \to 0$.

Proof. Note
$$\{x : \exists n,\, |f_n(x)| > g(x)\} \cup \{x : f_n(x) \not\to f(x)\}$$
has measure 0 because it is the countable union of measure 0 sets. Let $\hat{E}$ be the set of x such that $f_n(x) \to f(x)$ and $|f_n(x)| \leq g$. Then $\mu(E \backslash \hat{E}) = 0$, as mentioned above. Thus it suffices to show the theorem, without the "almost everywhere" business.

Now $|f| \le g$ gives $|f_n - f| \le 2g$. By the Fatou Lemma, (using $\left|\int h\right| \le \int |h|$)

$$\varlimsup_{n\to\infty} \int |f_n - f| \le \int \varlimsup_{n\to\infty} |f_n - f| = 0.$$

□

Note this is not true without the $|f_n(x)| \le g$ hypothesis. (Consider a "moving spike.")

Theorem 15.2: Suppose $\{f_n : n \ge 1\} \cup \{f\} \subseteq L^1(\mu; \mathbb{R})$, and that $f_n \to f$ (a.e., μ). Then

$$\lim_{n\to\infty} |\|f_n\|_{L^1} - \|f\|_{L^1} - \|f_n - f\|_{L^1}| = \lim_{n\to\infty} \int ||f_n| - |f| - |f_n - f|| \, d\mu = 0.$$

If $\|f_n\|_{L^1} \to \|f\|_{L^1} < \infty$, then $\|f - f_n\|_{L^1} \to 0$.

Proof. Note $||f_n| - |f| - |f_n - f|| \to 0$ (a.e., μ), and by the Triangle Inequality, $||f_n| - |f| - |f_n - f|| \le ||f_n| - |f_n - f|| + |f| \le 2|f|$. Thus the right-hand equality holds by Lebesgue's Dominated Convergence Theorem.

By the Triangle inequality,

$$|\|f_n\|_{L^1} - \|f\|_{L^1} - \|f_n - f\|_{L^1}| \le \int ||f_n| - |f| - |f_n - f|| \, d mu.$$

Taking $n \to \infty$ gives the left-hand equality. □

§2 Convergence in μ-measure

Definition 15.3: A sequence of measurable functions f_n converges to f **in μ-measure** if for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mu(|f_n - f| \ge \varepsilon) = 0.$$

By Markov's inequality, if $\|f_n - f\|_{L^1} \to 0$, then $f_n \to f$ in μ-measure. Using the triangle inequality, f_n can only converge to one function f.

Failure of Fatou: Just because $\|f_n\|_{L^1} = \int |f_n| \, d\mu \to 0$ does not mean $\lim_{n\to\infty} f_n = 0$. Defining $f_{2^m+l} = 1_{[2^{-m}l, 2^{-m}(l+1)]}$ (that is, characteristic functions covering smaller and smaller intervals in $[0,1]$), we see $\lim_{n\to\infty} f_n(x) = 1$ for every x (because there's infinitely many intervals in the sequence covering it), but $\lim_{n\to\infty} \int_{[0,1]} f_n \, \lambda_{\mathbb{R}} = 0$.

In Fatou's Lemma, Lebesgue's Dominated Convergence Theorem, and Theorem 15.2, $f_n \to f$ almost everywhere can be replaced with $f_n \to f$ in μ-measure.

Note when $|E| = n$ with the discrete topology, $L^1(\mu; E)$ is homeomorphic to $\mathbb{R}^n$ via the natural identification.

Lecture 16
Wed. 3/9/11

§1 Convergence

Lemma 16.1 (Cauchy criterion for measurable functions)**:** Let $\{f_n : n \geq 1\}$ be (measurable) functions on $(E, \mathcal{B}, \mu)$. Then there exists f such that

$$\lim_{n\to\infty} \mu(\sup_{n>m} |f_n - f| \geq \varepsilon) = 0 \text{ for all } \varepsilon > 0 \tag{3}$$

if and only if

$$\lim_{m\to\infty} \mu(\sup_{n>m} |f_n - f_m| \geq \varepsilon) = 0 \text{ for all } \varepsilon > 0. \tag{4}$$

Moreover

1. (3) implies $f_n \to f$ almost everywhere and in μ-measure.
2. If $\mu(E) < \infty$ then (3) holds iff there exists f such that $f_n \to f$ almost everywhere.

Proof. For the " $\Longrightarrow$ " direction, use the Triangle Inequality to get

$$\mu(\sup_{n>m} |f_n - f_m| \geq \varepsilon) \leq \mu(\sup_{n>m} |f_n - f| \geq \frac{\varepsilon}{2}) + \mu_{n>m}(|f - f_m| \geq \frac{\varepsilon}{2}) \leq 2\mu(\sup_{n>m} |f - f_n| \geq \frac{\varepsilon}{2}).$$

(If the LHS condition holds one of the RHS conditions must hold.)

To prove the reverse direction, we need to show

$$\{x : \lim_{n\to\infty} f_n(x) \text{ does not exist in } \mathbb{R}\}$$

has measure 0; we write this as the union of a countable number of measure 0 sets. Using Cauchy's criterion, it equals

$$\bigcup_{l=1}^{\infty} \bigcap_{m=1}^{\infty} \bigcup_{n>m} \left\{ x : |f_n(x) - f_m(x)| \geq \frac{1}{l} \right\}.$$

(The negation of Cauchy's criterion, is that for some $\varepsilon = \frac{1}{l} > 0$, and all m, there exists $n > m$ such that $|f_n(x) - f_m(x)| > \varepsilon$.) From (4), we know that

$$\mu\left(\bigcap_{m=1}^{\infty} \bigcup_{n>m} \left\{ x : |f_n(x) - f_m(x)| \geq \frac{1}{l} \right\} \right) = 0 \tag{5}$$

because the intersection is a smaller set than any one of the sets

$$\{\sup_{n>m} |f_n - f_m| \geq \varepsilon\}$$

which have measure tending to 0 as $\varepsilon \to 0$. Hence (5) has measure 0. Defining $f = \lim_{n\to\infty} f_n$ where the limit is defined gives the desired function.

Item 1 is clear. Since $\mu(E) = 0$ then using the convergence of measure of a decreasing sequence of sets, Proposition 5.5(5), gives the second item. (Details left to reader!) $\square$

Now we have to bring the sup's outside.

Lemma 16.2 (Cauchy's criterion, for convergence of measure)**:** Suppose

$$\lim_{m\to\infty} \sup_{n>m} (\mu(|f_n - f_m|) \geq \varepsilon) = 0 \text{ for all } \varepsilon > 0.$$

Then there exists a subsequence $\{f_{n_\gamma} : \gamma \geq 1\}$ such that

$$\lim_{i\to\infty} \mu(\sup_{j>i} |f_{n_j} - f_{n_i}| \geq \varepsilon) = 0 \text{ for all } \varepsilon > 0,$$

and there exists f so that $f_n \to f$ in μ-measure. (Conversely, if $f_n \to f$ in μ-measure, then f satisfies the Cauchy Criterion and so there exists a subsequence as above.)

I.e. if we have Cauchy's criterion for convergence of measure, then we have a subsequence converging almost everywhere and in μ-measure. This will imply the original sequence converges in measure to the same thing.

Proof. Choose n_i increasing so that

$$\sup_{n>n_i} \mu(|f_n - f_{n_i}| \geq 2^{-i-1}) \leq 2^{-i-1},$$

by using the Cauchy criterion. (Note the discepancy in measure and the size of the set where the discrepancy occurs both form convergent series.)

Now

$$\mu(|f_{n_{i+1}} - f_{n_i}| \geq 2^{-i-1}) \leq 2^{-i-1}$$

This gives the following. (Write $f_{n_j} - f_{n_i}$ as a telescoping series. If the LHS expression is at least 2^{-i}, then at least on the the inequalities on the RHS must hold. This argument is key!)

$$\mu(\sup_{j>i} |f_{n_j} - f_{n_i}| \geq 2^{-i}) \leq \sum_{k=i}^{j-1} \mu(|f_{n_{k+1}} - f_{n_k}| \geq 2^{-k-1}) \leq 2^{-i}.$$

We've found a subsequence satisfying the desired condition; by Lemma 16.1, $f_{n_i} \to f$ almost everywhere and in μ-measure. To get a $f_n \to f$ in μ-measure we use the Triangle Inequality:

$$\mu(|f_n - f| \geq \varepsilon) \leq \mu(|f_n - f_{n_i}| \geq \frac{\varepsilon}{2}) + \mu(|f_{n_i} - f| \geq \frac{\varepsilon}{2}) \to 0.$$

(If a bunch of guys converge to something and we have Cauchy's criterion, everyone gets drawn along.) □

Theorem 16.3 (Fatou, v.2)**:** If $f_n \geq 0$ almost everywhere and $f_n \to f$ **in μ-measure** then

$$\int f\, d\mu \leq \varliminf_{n\to\infty} \int f_n\, d\mu.$$

Proof. It suffices to consider $f_n \geq 0$ everywhere. Choose a subsequence $\{f_{n_m} : m \geq 1\}$ such that

$$\lim_{m\to\infty} \int f_{n_m}\, d\mu = \underline{\lim}_{n\to\infty} \int f_n\, d\mu.$$

Now by Theorem 16.2 choose a subsequence $\{f_{n_{m_i}} : i \geq 1\}$ such that $f_{n_{m_i}} \to f$ almost everywhere and apply Fatou's Lemma (Theorem 14.5) to this subsequence. $\square$

Theorem 16.4: $L^1(\mu)$ is a complete metric space. In other words, if

$$\sup_{n>m} \|f_n - f_m\|_{L^1} \to 0 \text{ as } n \to \infty$$

then there exists $f \in L^1$ such that $\|f_n - f\|_{L^1} \to 0$ as $n \to \infty$.

Proof. By Markov's inequality

$$\mu(|f_n - f_m| \geq \varepsilon) \leq \frac{1}{\varepsilon} \|f_n - f_m\|_{L^1}.$$

Thus by Theorem 16.2 there exist f such that $f_n \to f$ in μ-measure.

Now we show $f_n \to f$ in the L^1 norm. By Triangle,

$$\|f_n - f\|_{L^1} \leq \|f_n - f_m\|_{L^1} + \|f_m - f\|_{L^1} \to 0 :$$

The first goes to 0 as $n \to \infty$ by hypothesis. By Fatou v.2, since $f_m - f_n \to f_m - f$,

$$\|f_m - f\|_{L^1} \leq \underline{\lim}_{n\to\infty} \|f_m - f_n\|_{L^1} \leq \sup_{n \geq m} \|f_m - f_n\|_{L^1} \to 0$$

by our Cauchy criterion. $\square$

Now we prove separability (existence of countable dense subset). Let $(E, \mathcal{B}, \mu)$ be a finite measure space and suppose $\mathcal{B} = \sigma(\mathcal{C})$ where $\mathcal{C}$ is a Π-space, and $E \in \mathcal{C}$. For example, given have a separable metric space, we can find a countable number of balls so every open set is a union of some of the balls.

Theorem 16.5 (Dense subset in L^1)**:** Suppose μ is a finite measure. Let S be the set of all functions

$$\sum_{m=1}^{n} q_m 1_{C_m}, \quad C_m \in \mathcal{C}$$

Then S is dense in $L^1(\mu)$, i.e.

$$\overline{S} = L^1(\mu).$$

Thus if $\mathcal{C}$ is countable then $L^1(\mu)$ is separable.

Proof. It's easy to see that $\overline{S}$ must be a subspace (since the integral is linear), so it suffices to show all nonnegative measurable functions are in $\overline{S}$.

Every nonnegative function in $L^1(\mu)$ can be written as an increasing sequence of simple functions (see construction of Lebesgue integral); these simple functions approach it in μ-measure. Hence it suffices to show that simple functions are in $\overline{S}$, and by our note above, it suffices to show $1_\Gamma \in \overline{S}^{L^1(\mu)}$ for measurable Γ. But the class of Γ such that $1_\Gamma \in \overline{S}$ is a Λ-system and Π-system, so by Lemma 6.3, it is a σ-algebra, and hence contains $\mathcal{B}$, as needed. □

We can extend this result to σ-finite measures.

Definition 16.6: μ is σ-**finite** if there exists a set $\{E_m : n \ge 1\} \subseteq \mathcal{B}$ with

$$E = \sum_{n=1}^{\infty}, \quad \mu(E_n) < \infty.$$

(We can take them to be increasing or disjoint, if we wished.)

Theorem 16.7: Let E be a metric space. Let E_n be open sets increasing to E with $\mu(E_n) < \infty$. (σ-finite metric space where sets can be chosen to be open) Let $\mathcal{U}_n$ be the set of bounded uniformly continuous functions vanishing off of E_n. Then

$$\overline{\bigcup_{n=1}^{\infty} \mathcal{U}_n} = L^1(\mu).$$

Proof. Similar argument. Extra step saying you can approximate the indicator function of open set in terms of uniform functions. □

Lecture 17
Fri. 3/11/11

§1 Sunrise Lemma

Definition 17.1: A function g is **upper semicontinuous** if

$$g(x) \ge \overline{\lim_{y\to\infty}}\, g(y).$$

This is equivalent to $\{x : g(x) < a\}$ being open for all a.

Lemma 17.2 (Sunrise lemma)**:** Suppose $g : \mathbb{R} \to \mathbb{R}$ is upper semicontinuous and right continuous, and that $\lim_{x\to\pm\infty} g(x) = \mp\infty$. Set

$$G = \{x : \exists y > x, g(y) > g(x)\}.$$

Then G is open, and any nonempty connected open component of G is a bounded interval (a, b) such that $G(b) = G(a)$.

Think of g as mountainous terrain, going up to infinity at the left and down infinitely at the right. The set G is that portion of the terrain which will be in shadow the instant the sun comes over the horizon infinitely far at the right—those points where there's something to the right blocking the sun. The Sunrise Lemma says if you look at region in shadow, then is valley surrounded by hills with equal height.

Proof. Openness follows by upper continuity.

Take $c \in (a, b)$. Let
$$x = \sup\{x \ge c : g(x) \ge g(c)\},$$
the point beyond which nothing is higher than x. Now $g(x) \ge g(c)$ by upper semicontinuity. (Take y's in the set converging to x. Note x is finite since $\lim_{n\to\infty} g(x) = -\infty$.) Then $x \notin G$, because any point to the right of x higher than $g(x)$ would also be greater than $g(c)$, contradicting maximality. This means $b \le x$.

Also $g(c) \le g(b)$. (Draw a picture!)

Now $g(x) \le g(b)$ for $x \in (a, b)$. In particular $a > -\infty$. By right continuity, $g(a) \le g(b)$. But if $g(a) < g(b)$, then $b > a$ while $g(b) > g(a)$, and actually $a \in G$, contradiction. Hence $g(a) = g(b)$. □

§2 Lebesgue's differentiation theorem

Definition 17.3: Define the left, right, and two-sided **Lipschitz constants** of F at x to be

$$\begin{aligned}
\mathcal{L}_-F(x) &= \sup_{h<0} \frac{F(x+h)-F(x)}{h} \\
\mathcal{L}_+F(x) &= \sup_{h>0} \frac{F(x+h)-F(x)}{h} \\
\mathcal{L}F(x) &= \sup_{|h|>0} \frac{F(x+h)-F(x)}{h} = \max(\mathcal{L}_+F(x), \mathcal{L}_-F(x)).
\end{aligned}$$

Theorem 17.4: Let $F : \mathbb{R} \to \mathbb{R}$ be a nondecreasing function. For almost all $x \in \mathbb{R}$ (with respect to $\lambda_{\mathbb{R}}$),
$$F'(x) = \lim_{h\to 0} \frac{F(x+h)-F(x)}{h}$$
exists in $\mathbb{R}$.

We can replace monotonicity by bounded variation (since a function of bounded variation is the difference of two monotone functions). It suffices to prove this in an interval. Thus we can assume F is bounded, and by shifting by a constant, $F(-\infty) = \lim_{x\to-\infty} F(x) = 0$. A monotonous function is discontinuous at a countable number of points (a set of measure 0), so by redefining F at these points we may assume F is right continuous. The main steps are as follows. We assume F has the properties above.

1. The set of points at which F has large Lipschitz constant is small (Lemma 17.5 and Corollary 17.6). This is proved with the Sunrise Lemma (17.2).

2. If F is a definite integral $\int_{-\infty}^{x} f(y)\,\lambda_{\mathbb{R}}(dy)$ then F' exists and equals f almost everywhere (Theorem 17.7).

3. If F is Lipschitz then F is a definite integral (Lemma 18.1).

4. Take the Lipschitz constant to ∞ to conclude, if F is absolutely continuous then F is a definite integral (Lemma 18.2), and hence differentiable a.e. by item 2.

5. If F is singular then F is differentiable a.e., in fact with $F' = 0$.

6. Each F can be written as the sum of a absolutely continuous function and a singular function. From items 4 and 5, F is differentiable a.e.

Lemma 17.5:

$$\lambda_{\mathbb{R}}(\{x : \mathcal{L}F(x) > R\}) \leq \frac{2F(\infty)}{R}.$$

Proof. Consider $g(x) = F(x) - Rx$. It is right continuous and upper semicontinuous (g is increasing so we have upper semicontinuity on the left). Let G be as in the Sunrise Lemma (17.2). Then

$$G = \{x : \mathcal{L}_+F(x) > R\}.$$

Write G as union of disjoint components: $G = \bigcup_{n\geq 1}(a_n, b_n)$. By the Sunrise Lemma, $G(a_n) = G(b_n)$, giving $F(b_n) - F(a_n) = R(b_n - a_n)$. Now

$$\lambda_{\mathbb{R}}(G) = \sum_{n=1}^{\infty}(b_n - a_n) = \frac{1}{R}\sum_{n=1}^{\infty}(F(b_n) - F(a_n)) \leq \frac{F(\infty)}{R}.$$

Similarly, $\lambda_{\mathbb{R}}(\{x : \mathcal{L}_+F(x) > R\}) \leq \frac{F(\infty)}{R}$ and the result follows. □

Corollary 17.6: $\lambda_{\mathbb{R}}(\mathcal{L}F = \infty) = 0$.

We next show a generalization of the fundamental theorem of calculus.

Theorem 17.7: Suppose that $F(x) = \int_{(-\infty,x]} f(y)\,\lambda_{\mathbb{R}}(dy)$ where $f \geq 0$ and $f \in L^1$. Then $F'(x)$ exists and equals $f(x)$ $\lambda_{\mathbb{R}}$-almost everywhere.

Proof. Let $\mathcal{G}$ be the set of all $f \in L^1(\lambda_{\mathbb{R}})$ such that

$$\lim_{I\searrow\{x\}} \frac{1}{|I|}\int_I |f(y) - f(x)|\lambda_{\mathbb{R}}(dy) = 0 \text{ almost everywhere.}$$

(I represents an interval.) Note in any case the LHS is at least

$$\lim_{I\searrow\{x\}} \left|\frac{1}{|I|}\int_I f - f(x)\right|. \tag{6}$$

By Theorem 16.7 the set of continuous functions with compact support in L^1 is dense in L^1. Since continuous functions are in $\mathcal{G}$, it suffices to show $\mathcal{G}$ is closed under L^1-convergence.

Let $f_n \in \mathcal{G}$ with and $f_n \to f$ in L'. Then

$$\begin{aligned}\lambda_{\mathbb{R}}\left(\left\{x : \overline{\lim_{I\searrow\{x\}}} \frac{1}{|I|}\int_I |f-f(x)|\,d\lambda_{\mathbb{R}} \geq \varepsilon\right\}\right) &\leq \lambda_{\mathbb{R}}\left(\left\{x : \overline{\lim_{I\searrow\{x\}}} \frac{1}{|I|}\int_I |f-f_n|\,d\lambda_{\mathbb{R}} \geq \frac{\varepsilon}{3}\right\}\right)\\ &+ \lambda_{\mathbb{R}}\left(\left\{x : \overline{\lim_{I\searrow\{x\}}} \frac{1}{|I|}\int_I |f_n-f_n(x)|\,d\lambda_{\mathbb{R}} \geq \frac{\varepsilon}{3}\right\}\right)\\ &+ \lambda_{\mathbb{R}}(|f_n - f| \geq 1).\end{aligned}$$

By (17.5) and (6), the first and third terms are at most $\frac{2\|f-f_n\|_{L^1}}{\varepsilon}$ and $\frac{\|f-f_n\|_{L^1}}{\varepsilon}$ and by assumption the second term is 0, so the RHS is at most $\frac{\|f-f_n\|_{L^1}}{\varepsilon}$, as needed.

□

For $f \in L^1$, define the **Hardy-Littlewood maximal function** of f to be

$$Mf(x) = \sup_{I \ni \{x\}} \frac{1}{|I|}\int_I |f(y)|\lambda_R(dy)$$

Now by (17.5),

$$\lambda_{\mathbb{R}}(Mf > R) \leq \frac{2\,\|f\|_{L^1}}{R}.$$

(This is the Hardy-Littlewood inequality.)

Lecture 18
Mon. 3/14/11

§1 Lebesgue's differentiation theorem

Lemma 18.1: Suppose $F(-\infty) = 0$ and F satisfies the Lipschitz condition

$$0 \leq F(y) - F(x) \leq L(y-x) \text{ for } x \leq y.$$

Then F is a definite integral.

Proof. Let

$$f_n(x) = 2^n(F((k+1)2^{-n}) - F(k2^{-n})),\; k2^{-n} \leq x < (k+1)2^{-n}.$$

Then $0 \leq f_n \leq L$ and

$$\int_{-\infty}^{\infty} f_n\,dx = F(\infty).$$

Let $m < n$. Consider

$$\begin{aligned}\int_{[k2^{-m},(k+1)2^{-m})} f_m f_n \, d\lambda &= f_m(k2^{-m}) \int f_n \, d\lambda \\ &= 2^{-m} f_m(k2^{-n})^2 \\ &= \int f_m^2 \, d\lambda.\end{aligned}$$

Hence

$$\int (f_n - f_m)^2 \, d\lambda = \int f_n^2 \, d\lambda - \int f_m^2 \, d\lambda \to 0$$

since $\int f_n^2 \leq LF(\infty)$. (A bounded monotonic sequence converges.) By Markov's inequality,

$$\mu(|f_n - f_m| \geq \varepsilon) = \mu(|f_n - f_m|^2 \geq \varepsilon^2) \leq \frac{1}{\varepsilon^2} \int |f_n - f_m|^2 \, d\lambda.$$

Taking the sup over $n > m$,

$$\sup_{n>m} \mu(|f_n - f_m| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \sup_{n>m} \int |f_n - f_m|^2 \to 0.$$

By Cauchy's Criterion for convergence in measure (16.2), there exists f so that $f_n \xrightarrow{\mu} f$. Changing f on a set of measure 0 we may assume $0 \leq f \leq L$.

Consider $F(y) - F(x)$. For $n \geq m$,

$$F(l2^m) - F(k2^{-m}) = \int_{[k2^{-m},l2^{-m})} f_n \, d\lambda \to \int f \, d\lambda \text{ as } n \to \infty.$$

by Lebesgue's Dominated Convergence Theorem for convergence in measure (L is a Lebesgue dominant of the f_m). Let $F(y) = \int_{(-\infty,y]} f$. Hence for dyadics,

$$F(y) - F(x) = \int_{(x,y]} f \, d\lambda.$$

Both sides are continuous and the dyadics for a dense set so this is true for all x and y. $\square$

Lemma 18.2: Suppose F is absolutely continuous and nondecreasing. Then F is a definite integral. (Hence, by Theorem 17.7, it is differentiable a.e.)

Proof. For every $\varepsilon > 0$ and $\delta > 0$ there exist intervals I such that $\sum |I| < \delta$ and

$$\sum_I F(I^+) - F(I^-) < \varepsilon.$$

From a homework problem, F is absolutely continuous iff μ_F is absolutely continuous with respect to λ_R (i.e. $\lambda(\Gamma) = 0 \implies \mu(\Gamma) = 0$).

By Corollary 17.6, $\mu_F(\mathcal{L}F = \infty) = 0$. Let

$$F_n(x) = \mu_F((-\infty, x] \cap \{\mathcal{L}F \le n\}).$$

It is nondecreasing, tends to 0 a $-\infty$, and increases to $F(x)$ as $n \to \infty$, because we've remove a set whose measure tends to 0. (The measure tends to $\mu((-\infty, x] \cap \{\mathcal{L}F < \infty\}) = 0$.)

We claim that F_n is Lipschitz with constant n: $F_n(y) - F_n(x) \le n(y - x)$. Indeed, suppose $c \in (x, y] \cap \{\mathcal{L}F \le n\}$. If $(x, y]$ does not intersect $\{\mathcal{L}F \le n\}$ then the difference is 0. Else, choose c so that $c \in (x, y] \cap \{\mathcal{L}F \le n\}$. Then

$$\begin{aligned} F_n(y) - F_n(x) &= \mu_F((x, y] \cap \{\mathcal{L}F \le n\}) \\ &\le (F(y) - F(x)) \\ &= (F(y) - F(x)) - (F(c) - F(x)) \\ &\le n(y - x). \end{aligned}$$

Then by Lemma 18.1,

$$\begin{aligned} F_n(x) &= \int_{(-\infty,x]} f_n \\ F_{n+1}(x) - F_n(x) &= \int_{(-\infty,x]} f_{n+1} - f_n \, d\lambda \\ &= \mu_F((-\infty, x] \cap \{n < \mathcal{L}F \le n + 1\}). \end{aligned}$$

Hence F_n is nonnegative nonincreasing, and by the Monotone Convergence Theorem,

$$F(x) = \lim_{n\to\infty} F_n(X) = \lim_{n\to\infty} \int_{(-\infty,x]} f_n \, d\lambda = \int_{(-\infty,x]} f \, d\lambda_{\mathbb{R}}.$$

$\square$

Theorem 18.3: Let F be a bounded, right-continuous, non-decreasing function with $F(-\infty) = 0$.

Then F can be decomposed into an absolutely continuous, non-decreasing F_a and a singular, right-continuous, non-decreasing F_b: $F = F_a + F_b$.

Proof. Define the measures

$$\begin{aligned} \mu_a(\Gamma) &= \mu_F(\Gamma \cap \{\mathcal{L}F < \infty\}) \\ \mu_b(\Gamma) &= \mu_F(\Gamma \cap \{\mathcal{L}F = \infty\}) \end{aligned}$$

and let

$$\begin{aligned} \mathcal{F}_a(x) &= \mu_a((0, x]) \\ \mathcal{F}_b(x) &= \mu_b((0, x]) \end{aligned}$$

be the corresponding distribution functions. Since $\lambda_{\mathbb{R}}(\mathcal{L}F = \infty) = 0$, μ_b is singular to $\lambda_{\mathbb{R}}$, so by Exercise 2.2.39, the distribution function F_b is singular.

Since $\mu_a(\mathcal{L}F_a = \infty) \le \mu_F(\mathcal{L}F = \infty) = 0$, F is absolutely continuous by Exercise 2.2.38. $\square$

Lemma 18.4: If F is singular, then F' exists and equals 0 Lebesgue almost everywhere.

Proof. Since F is singular, μ_F is singular to $\lambda_{\mathbb{R}}$ by Exercise 2.2.38. Thus there exists a set B such that $\mu_F(B) = 0$ and $\lambda_{\mathbb{R}}(B^c) = 0$. Let $\{K_n\}$ be closed subsets of B^c such that $K_n \subseteq K_{n+1}$ and $\mu_F(B^c \backslash K_n) \le \frac{1}{n}$ for each $n \in \mathbb{N}$. Set

$$F_n(x) = \mu_F((-\infty, x]) \cap K_n.$$

For $x \in B$, x is at positive distance away from the closed set K_n, so $F_n'(x) = 0$.

For $x \in B$, $F_n'(x) = 0$ for each n, so this is true for F as well. Now

$$\begin{aligned}\lim_{n\to\infty} F_n(x) &= \mu_F((-\infty, x] \cap B^c). \\ &= \mu_F((-\infty, x]) && (B \text{ has measure } 0) \\ &= F(x).\end{aligned}$$

□

Any F satisfying the conditions of Lebesgue's differentiation theorem can be written as $F_a + F_b$ as in Theorem 18.3. Both F_a and F_b' have derivatives except on a set of measure 0, so the theorem follows.

Lecture 19
Wed. 3/16/11

§1 Product measures

Let $(E_1, \mathcal{B}_1)$ and $(E_2, \mathcal{B}_2)$ be measure spaces. The product of the σ-algebras is defined to be the smallest σ-algebra generated by the products of measurable sets:

$$\mathcal{B}_1 \times \mathcal{B}_2 := \sigma(\{\Gamma_1 \times \Gamma_2\} : \Gamma_i \in \mathcal{B}_i\}).$$

This is a Π-system. If μ and ν are finite measures on $(E_1 \times E_2, \mathcal{B}_1 \times \mathcal{B}_2)$, and if $\mu(\Gamma_1 \times \Gamma_2) = \nu(\Gamma_1 \times \Gamma_2)$ for all $\Gamma_i \in \mathcal{B}_i$, then $\mu = \nu$. (Theorem 6.2, if two measures agree on a Π-system they agree on the σ-algebra generated by the Π-system.)

If E_1, E_2 are topological spaces then

$$\mathcal{B}_{E_1} \times \mathcal{B}_{E_2} \subseteq \mathcal{B}_{E_1 \times E_2}.$$

(If G_1, G_2 are open sets in E_1, E_2 then $G_1 \times G_2$ is open in $E_1 \times E_2$ so in $\mathcal{B}_{E_1 \times E_2}$.) Equality does not hold in general, but holds when E_1 and E_2 are second countable: In this case every open set of $E_1 \times E_2$ is a *countable* union of products of open sets in E_1 and E_2.

We would like to produce a measure $\mu_1 \times \mu_2$ on $(E_1 \times E_2, \mathcal{B}_1 \times \mathcal{B}_2)$ such that

$$(\mu_1 \times \mu_2)(\Gamma_1 \times \Gamma_2) = \mu_1(\Gamma_1) \times \mu_2(\Gamma_2).$$

Note we already know there is at most one such measure. Define

$$(\mu_1 \times \mu_2)(\Gamma) = \int_{E_1} \left(\int_{E_2} 1_\Gamma(x_1, x_2) \mu_2(dx_2) \mu_1(dx_1) \right). \tag{7}$$

We know this gives the correct value for $\Gamma = \Gamma_1 \times \Gamma_2$, and it is a measure if well-defined. But we need to show we can carry out the integrals, i.e. $1_\Gamma(x_1, \cdot)$ is measurable.

Definition 19.1: A collection $\mathcal{L}$ of functions $f : E \to (-\infty, \infty]$ is a semi-lattice if $f^\pm \in \mathcal{L}$ when $f \in \mathcal{L}$. $\mathcal{K} \subseteq \mathcal{L}$ is an $\mathcal{L}$-system if it has the following properties.

1. $1 \in \mathcal{K}$.
2. If $f, g \in \mathcal{K}$ and $f \leq g$, then $g - f \in \mathcal{K}$ if $g - f \in \mathcal{L}$.
3. If $f, g \in \mathcal{K}$ and $\alpha, \beta \geq 0$, then $\alpha f + \beta g \in \mathcal{K}$.
4. If $\{f_n\}$ is a sequence of functions in $\mathcal{K}$ and $f_n \nearrow f$, then $f \in \mathcal{K}$ if $f \in \mathcal{L}$ or if f is bounded.

Lemma 19.2: Let $(E, \mathcal{B})$ be a measure space with $\mathcal{B} = \sigma(\mathcal{C})$ where $\mathcal{C}$ is a Π-system. If $\mathcal{K}$ is a $\mathcal{L}$-system of functions $f : E \to (-\infty, \infty]$ and if $1_\Gamma \in \mathcal{K}$ for $\Gamma \in \mathcal{C}$, then $\mathcal{K}$ contains all $\mathcal{B}$-measurable $f \in \mathcal{L}$.

This the function analogue of the fact that the smallest Λ-system containing a Π-system contains the minimal σ-algebra generated by the Π-system. (See proof of Theorem 6.2.)

Proof. The indicator functions of $\Gamma \in \mathcal{B}$ are in $\mathcal{K}$ by the fact cited above. Hence all finite positive linear combinations of them—nonnegative simple functions, are in $\mathcal{K}$. Now every nonnegative measurable function (in $\mathcal{L}$) is the limit of an increasing sequence of simple functions, so is in $\mathcal{K}$. For general f decompose into positive and negative parts. □

Lemma 19.3: Suppose $f(x_1, x_2)$ is measurable with respect to $\mathcal{B}_1 \times \mathcal{B}_2$. For every $x_1 \in E_1$, $f(x_1, \cdot)$ is measurable on $(E_2, \mathcal{B}_2)$, and

$$\int_{E_1} f(x_1, \cdot)\, \mu_1(dx_1)$$

is measurable on $(E_2, \mathcal{B}_2)$.

Proof. By the Monotone Convergence Theorem, we only need to check for bounded f. Let $\mathcal{L}$ be the collection of bounded functions on $E_1 \times E_2$ and let $\mathcal{K}$ the subset of $\mathcal{L}$ which have the properties above. Then $\mathcal{K}$ is a $\mathcal{L}$-system so by Lemma 19.2, $\mathcal{K} = \sigma(\mathcal{C})$, where $\mathcal{C} = \{\Gamma_1 \times \Gamma_2 : \Gamma_i \in \mathcal{B}_i\}$. □

Theorem 19.4: Given finite measure spaces $(E_1, \mathcal{B}_1, \mu_1)$ and $(E_2, \mathcal{B}_2, \mu_2)$, there is a unique measure on $(E_1 \times E_2, \mathcal{B}_1 \times \mathcal{B}_2)$ such that

$$\nu(\Gamma_1 \times \Gamma_2) = \mu_1(\Gamma_1)\mu_2(\Gamma_2)$$

whenever $\Gamma_i \in \mathcal{B}_i$. It equals

$$\int_{E_1}\int_{E_2} 1_\Gamma(x_1, x_2)\, \mu_2(dx_2)\, \mu_1(dx_1) = \int_{E_2}\int_{E_1} 1_\Gamma(x_1, x_2)\, \mu_1(dx_1)\, \mu_2(dx_2)$$

Proof. This follows directly by letting ν be (7) and using Lemma 19.3. Equality above follows from uniqueness. □

Theorem 19.5 (Tonelli)**:** Let $(E_1, \mathcal{B}_1, \mu_1)$ and $(E_2, \mathcal{B}_2, \mu_2)$ be σ-finite measure spaces (i.e. they are countable union of measurable sets with finite measure). Then the construction in Theorem 19.4 still works.

Proof. Choose pairwise disjoint $\{E_{i,n} : n \geq 1\} \subseteq B$ such that $E_i = \bigcup_{n=1}^\infty E_{i,n}$. Define $\mu_{i,n}(\Gamma_i) = \mu_i(\Gamma_i \cap E_{i,n})$, $\Gamma_i \cap \mathcal{B}_i$, and let $\nu_{m,n}$ be the measure constructed in 19.4 from $\mu_{1,m}$ and $\mu_{2,n}$. The desired measure is

$$(\mu_1 \times \mu_2)(\Gamma) = \sum_{m,n \geq 1} \nu_{m,n}(\Gamma).$$

This can be written in terms of integrals, by noting

$$\int_{E_2} f(\cdot, x_n)\, \mu_2(dx_2) = \sum_{n=1}^\infty \int_{E_{2,n}} f(\cdot, x_2)\, \mu_{2,n}(dx_2).$$

□

Theorem 19.6 (Fubini)**:** Given σ-finite measure spaces, f is $\mu_1 \times \mu_2$-integrable iff

$$\int_{E_1} \left(\int_{E_2} |f(x_1, x_2)|\, \mu_2(dx_2) \right) \mu_1(dx_1) < \infty$$

iff

$$\int_{E_2} \left(\int_{E_1} |f(x_1, x_2)|\, \mu_1(dx_1) \right) \mu_2(dx_2) < \infty.$$

If f is $\mu_1 \times \mu_2$-integrable, then

$$\int_{E_1}\int_{E_2} f(x_1, x_2)\, \mu_2(dx_2)\, \mu_1(dx_1)$$

equals both integrals above. (The inner integral is ∞ on a set of measure 0.)

Proof. This is true for characteristic functions. Use the argument in Lemma 19.3, namely use Lemma 19.2 to show

$$\int_{E_1}\int_{E_2} g(x_1,x_2)\,\mu_2(dx_2)\,\mu_1(dx_1) = \int_{E_2}\int_{E_1} g(x_1,x_2)\,\mu_1(dx_1)\,\mu_2(dx_2).$$

for nonnegative $\mu_1\times\mu_2$-measurable g when the measure is finite; use the argument in Tonelli to extend it to when it is σ-finite. Put in $g=|f|$ to show the first part.

To show the second part split up the positive and negative part of f. □

Lecture 20
Fri. 3/18/11

§1 Isodiametric inequality

Theorem 20.1 (Isodiametric inequality)**:**

$$\lambda_{\mathbb{R}^N}(\Gamma) \le \Omega_N \operatorname{rad}(\Gamma)^N.$$

Proof. If $-x\in\Gamma$ whenever $x\in\Gamma$, then $\Gamma\subseteq B(0,\operatorname{rad}(\Gamma))$. Indeed, $x\in\Gamma$ implies $2|x| = |x-(-x)| \le 2\operatorname{rad}(\Gamma)$. Thus the theorem holds when Γ is symmetric.

Now we use the technique of Steiner Symmetrization to reduce the general case to the symmetric case.

Theorem 20.2 (Steiner symmetrization)**:** Let e be a unit vector in S^{N-1}. Let

$$P(e)=\{x\in\mathbb{R}^N : x\perp e\}.$$

Given a point $\xi\in P(e)$, look at the straight line passing through ξ in the direction e, and consider the intersection of it with Γ. Let $l(\Gamma,e,\xi)$ be the one-dimensional Lebesgue measure of that intersection:

$$l(\Gamma,e,\xi)=\lambda_{\mathbb{R}}(\{t\in\mathbb{R}:\xi+te\in\Gamma\}).$$

Define the symmetrization of Γ with respect to e by

$$S(\Gamma,e)=\{\xi+te : |t|<\frac{1}{2}l(\Gamma,e,\xi)\}.$$

(Think of the set as gooey stuff. Poke a stick through it in the direction e and squoosh the gooey stuff to remove the holes and so that half of the stuff is on each side of $P(e)$.)

Then

$$\operatorname{rad}(S(\Gamma,e)) \le \operatorname{rad}(\Gamma) \tag{8}$$
$$\lambda_{\mathbb{R}_N}(S(\Gamma,e)) = \lambda_{\mathbb{R}^N}(\Gamma). \tag{9}$$

Furthermore, if $T_{\mathcal{O}}$ is an orthogonal transformation such that $T_{\mathcal{O}}\Gamma=\Gamma$ and $T_{\mathcal{O}}e=\pm e$ then $S(\Gamma,e)=T_{\mathcal{O}}S(\Gamma,e)$. The symmetrized object will be invariant under the orthogonal transformation.

Proof. We may assume $e = (0, \dots, 0, 1)$; then $P(e) = \mathbb{R}^{N-1} \times \{0\}$. Let

$$f(\xi) = l(\Gamma, e, (\xi, 0)) = \int 1_\Gamma((\xi, t))\, dt.$$

(Note this is measurable by Lemma 19.3.) Now by Fubini's Theorem (19.6),

$$\lambda_{\mathbb{R}}(\Gamma) = \int_{\mathbb{R}^{N-1}} f(\xi)\, d\xi = \int_{\mathbb{R}^{N-1}} \int_{\mathbb{R}} 1_\Gamma((\xi, t))\, dt\, d\lambda_{\mathbb{R}^{N-1}}.$$

For the second part, we may assume Γ is bounded, and Γ is closed (since $\overline{\Gamma} = \Gamma$), hence compact. Suppose $x, y \in S(\Gamma, e)$ with $x = (\xi, s)$ and $y = (\eta, t)$. Let

$$\begin{aligned} M^+(x) &= \sup\{\alpha : (\xi, \alpha) \in \Gamma\} \\ M^+(y) &= \sup\{\alpha : (\eta, \alpha) \in \Gamma\} \\ M^-(x) &= \inf\{\alpha : (\xi, \alpha) \in \Gamma\} \\ M^-(y) &= \inf\{\alpha : (\eta, \alpha) \in \Gamma\} \end{aligned}$$

(The lowest and highest points in Γ along the direction e from x or y.) Note $M^+(x) - M^-(x) \ge 2|s|$. Similarly $M^+(y) - M^-(y) \ge 2|t|$. Now

$$\begin{aligned} \max(M^+(y) - M^-(x), M^+(x) - M^-(y)) &\ge \frac{M^+(y) - M^-(x)}{2} + \frac{M^+(x) - M^-(y)}{2} \\ &= \frac{M^+(x) - M^-(y)}{2} + \frac{M^+(y) - M^-(y)}{2} \\ &\ge |s| + |t|. \end{aligned}$$

Using the distance formula $|y - x|^2 \le 4 \operatorname{rad}(\Gamma)^2$.

The last assertion is clear. $\square$

Take an orthonormal basis $e_1, \dots, e_N$. Let $\Gamma_0 = \Gamma$ and $\Gamma_{n+1} = S(\Gamma_n, e_{n+1})$. We pick up more and more directions of symmetry until it's completely symmetric. (Γ_1 is invariant under reflection in the first coordinate, Γ_2 is invariant under orthogonal transformations changing the first and second coordinate, and so on.) Now

$$\lambda_{\mathbb{R}^N}(\Gamma) = \lambda_{\mathbb{R}^N}(\Gamma_N) \le \Omega_N \operatorname{rad}(\Gamma)^N$$

by the symmetric case, as needed. $\square$

§2 Hausdorff's construction

Definition 20.3: Given $\delta > 0$, define

$$H^{N,\delta}(\Gamma) = \inf\left\{ \sum_{C \in \mathcal{C}} \Omega_N \operatorname{rad}(C)^N : \Gamma \subseteq \bigcup \mathcal{C} \text{ and } \|\mathcal{C}\| \le \delta \right\}.$$

Note as a function of delta, this increases as delta decreases. The **Hausdorff measure** of Γ is

$$H^N(\Gamma) = \lim_{\delta \searrow 0} H^{N,\delta}(\Gamma).$$

Theorem 20.4 (Lebesgue and Hausdorff measure coincide for measurable sets)**:** If $\Gamma \in \mathcal{B}_{\mathbb{R}^N}$ then $H^N(\Gamma) = H^{N,\delta}(\Gamma) = \lambda_{\mathbb{R}^N}(\Gamma)$.

Proof. The inequality $\lambda_{\mathbb{R}^N}(\Gamma) \le H^{N,\delta}(\Gamma)$ holds by the Isodiametric Inequality. The opposite inequality uses the following.

Lemma 20.5 (Covering lemma)**:** Given G open, there exists a sequence $\{B_k : k \ge 1\}$ of closed mutually disjoint balls such that $G \supseteq \bigcup_{n=1}^{\infty} B_n$ and

$$\lambda_{\mathbb{R}^N}\left(G \setminus \bigcup_{k=1}^{\infty} B_k\right) = 0.$$

I.e. we can cover almost all of G with a countable number of disjoint balls.

Proof. Recall we can write G as a union of a countable number of nonoverlapping cubes $G = \bigcup_{n=1}^{\infty} Q_n$. Put a circle inside each cube; they occupy a volume that is a constant times the volume of the cube. Let these balls be $B_1, \ldots$; take enough of them (a finite number) so they've covered a fixed constant of the volume of G. Since we stopped with a finite number, $G \setminus \bigcup_{n=1}^{n_1} B_n$ is open. Repeat with this set and keep going. $\square$

Then $\lambda_{\mathbb{R}^N} \ge H^{N,\delta}(G)$. $\square$

Note we can also define $H^{s,\delta}$ for $s \ne N$:

$$H^{N,\delta}(\Gamma) = \inf\left\{\sum_{C \in \mathcal{C}} \Omega_s \operatorname{rad}(C)^s : \Gamma \subseteq \bigcup \mathcal{C} \text{ and } \|\mathcal{C}\| \le \delta\right\}.$$

Unlike before there is dependence of δ. Then the s-dimensional Hausdorff measure is

$$H^N(\Gamma) = \lim_{\delta \searrow 0} H^{N,\delta}(\Gamma).$$

For $s > 0$ the measure is identically 0. For $s < N$ it's not a finite measure, it's ∞ for open sets; it's meaningful for "s-dimensional" stuff. They provide generalization of surface measure.

Lecture 21
Mon. 3/28/11

§1 Change of variables

Let $(E_1, \mathcal{B}_1)$ and $(E_2, \mathcal{B}_2)$ be measure spaces and let $\Phi : E_1 \to E_2$ be a measurable map. Let μ be a measure on $(E_1, \mathcal{B}_1)$ and consider the measure $\Phi_* \mu = \mu \circ \Phi^{-1}$ on $(E_2, \mathcal{B}_2)$.

Theorem 21.1: For all nonnegative $\mathcal{B}_2$-measurable functions φ,

$$\int_{E_2} \varphi \, d(\Phi_*) = \int_{E_1} \varphi \circ \Phi \, d\mu.$$

Proof. This holds for characteristic functions:

$$\int_{E_2} 1_\Gamma \, d(\Phi_*\mu) = (\Phi_*\mu)(\Gamma) = \int_{E_1} 1_\Gamma \circ \Phi \, d\mu.$$

Hence it holds for simple functions, and by taking monotone limits, all nonnegative monotone functions. □

In order for this to be useful, we need a nice expression for $\Phi_*\mu$.

Theorem 21.2: Let $-\infty < a < b < \infty$ and μ and $\psi : [a,b] \to \mathbb{R}$ a right-continuous non-decreasing function. If $\varphi : [a,b] \to \mathbb{R}$ is bounded, then φ is Riemann-integrable with respect to ψ iff φ is continuous μ_ψ-almost everywhere. (Recall $\mu_\psi((x,y]) = \psi(y) - \psi(x)$.)

Proof. φ is ψ-Riemann integrable iff for every $\varepsilon > 0$ there exists $\delta > 0$ such that for every cover $\mathcal{C}$ with $\|\mathcal{C}\|$,

$$\sum_{\substack{I \in \mathcal{C}, \\ \sup_I \varphi - \inf_I \varphi \geq \varepsilon}} \Delta_I \psi < \varepsilon. \tag{10}$$

(Exercise.)

Suppose φ is ψ-Riemann integrable. For n, choose $\mathcal{C}_n$ such that $\|\mathcal{C}_n\| \leq \frac{1}{n}$ and ψ is continuous at each left hand endpoint I^-. (There are countably many discontinuities.) Let

$$\mathcal{C}_{m,n} = \{I \in \mathcal{C}_n : \sup_I \varphi - \inf_I \varphi \geq \frac{1}{m}\}.$$

Let

$$\Delta = \bigcup_{n=1}^\infty \{I^- : I \in \mathcal{C}_n\}.$$

Each element of Δ is a continuity point of ψ, so has measure 0 under μ_ψ. Since Δ is countable, this implies $\mu_\psi(\Delta) = 0$. By definition of continuity

$$\{x \in (a,b] \backslash \Delta : \varphi \text{ is not continuous at } x\} \subseteq \bigcup_{m=1}^\infty \bigcap_{n=1}^\infty \bigcup \mathcal{C}_{m,n}.$$

(There's some m so that nearby points vary at least $\frac{1}{m}$.) Now

$$\mu_\psi(\{x \in (a,b] : \varphi \text{ is not continuous at } x\} \backslash \Delta\}) \leq \sum_{m=1}^\infty \mu_\psi\left(\bigcup_{n=1}^\infty \mathcal{C}_{m,n}\right) = 0$$

since

$$\mu_\psi\left(\bigcup_{n=1}^{\infty}\mathcal{C}_{m,n}\right) \le \varliminf_{n\to\infty}\mu_\psi(\mathcal{C}_{m,n}) = \varliminf_{n\to\infty}\sum_{I\in\mathcal{C}_{m,n}}\Delta_I\psi = 0$$

by (10). The other direction is similar. □

Theorem 21.3: Let $(E,\mathcal{B},\mu)$ be a measure space and $f : E \to [0,\infty]$ be $\mathcal{B}$-measurable. Then $\mu(f>t)$ (as a function of t) is right-continuous and non-increasing.

Suppose $\varphi \in C([0,\infty];\mathbb{R}) \cap C^1((0,\infty);\mathbb{R})$, φ is non-decreasing, and $\varphi(0) = 0 < \varphi(t)$ for $t > 0$. Then

$$\int_E \varphi \circ f \, d\mu = \int_{(0,\infty)} \varphi'(t)\mu(f>t)\,dt.$$

This is used in probability (expectation), the continuous version of the following: For X taking values in $\mathbb{N}$, $\mathbb{E}(X) = \sum_{n=0}^{\infty} P(X>n)$. Indeed, by interchanging order of summation

$$E(X) = \sum_{n=0}^{\infty} nP(X=n) = \sum_{n=0}^{\infty}\sum_{m=1}^{n} P(X=n) = \sum_{m=1}^{\infty}\sum_{n\ge m} P(X=n).$$

We use the same idea, with integration by parts.

Proof. Note

$$\{f>s\} = \bigcup_{t>s}\{f>t\}$$

so $\{f>t\} \nearrow \{f>s\}$ when $t \searrow s$. Hence $\mu(\{f>t\}) \nearrow \mu(\{f>s\})$.

The integral on the RHS is the limit of integrals over $[a,b]$ as $a\to 0$, $b\to\infty$. Then the RHS is an ordinary Riemann integral. $\mu(f>t)$ has at most a countable number of discontinuities so the integrand is Riemann integrable. Evaluate the RHS as a Riemann integral. Integration by parts! □

§2 Polar coordinates

In $\mathbb{R}^N$ let

$$\mathbb{S}^{N-1} = \{x : |x| = 1\}.$$

We look for a measure $\lambda_{\mathbb{S}^{N-1}}$ on $\mathbb{S}^{N-1}$ such that

$$\int_{\mathbb{R}^N} f(x)\,\lambda_{\mathbb{R}^N}(dx) = \int_{(0,\infty)} r^{N-1}\left(\int_{\mathbb{S}^N} f(r\omega)\lambda_{\mathbb{S}^{N-1}}(d\omega)\right)dr.$$

Define $\Phi : \mathbb{R}^N\backslash\{0\} \to \mathbb{S}^{N-1}$ by $\Phi(x) = \frac{x}{|x|}$. Define

$$\lambda_{\mathbb{S}^{N-1}} = N\Phi_*\lambda_{B(0,1)\backslash\{0\}}$$

Lecture 22
Wed. 3/30/11

§1 Change of variables

Suppose G is an open set in $\mathbb{R}^N$, and $\Phi : G \to \mathbb{R}^N$ is C^1. Let $\Phi = \begin{bmatrix} \Phi_1 \\ \vdots \\ \Phi_n \end{bmatrix}$ and define

$$\begin{aligned} \frac{\partial \Phi}{\partial x} &= (\partial_{x_1}\Phi, \ldots, \partial_{x_N}\Phi) \\ J\Phi &= \left| \det \frac{\partial \Phi}{\partial x} \right|. \end{aligned}$$

(If $\Phi(x) = Ax$ then $J\Phi = |\det A|$.)

Theorem 22.1: If $\Phi \in C^2(G; \mathbb{R}^N)$ is 1-to-1 and if $J\Phi > 0$, then

$$\lambda_{\mathbb{R}^N}(\Phi(\Gamma)) = \int_\Gamma J\Phi \, d\lambda_{\mathbb{R}^N}.$$

For $f \geq 0$ is measurable on $\Phi(G)$,

$$\int_G f \circ \Phi \, J\Phi \, d\lambda_{\mathbb{R}} = \int_{\Phi(G)} f \, d\lambda_{\mathbb{R}^N}.$$

Proof. We use the inverse function theorem: Let $\Phi \in C^k(G; \mathbb{R}^N)$ with $J\Phi > 0$. For $p \in G$ there exists $B(p,r) \subseteq G$ such that $\Phi|_{B(p,r)}$ is 1-to-1; moreover $\Phi^{-1} \in C^k(\Phi(B(p,r)), \mathbb{R}^N)$ and $\left(\frac{\partial \Phi^{-1}}{\partial y} \circ \Phi\right) \frac{\partial \Phi}{\partial x} = I$. As a corollary, Φ takes open/closed/measurable subsets of G into open/closed/measurable sets. (Each of these is locally true as Φ is locally a diffeomorphism.)

We first show that

$$\int_\Phi f \, d\lambda \leq \int_G f \circ \Phi J\Phi \, d\lambda. \tag{11}$$

If Φ is 1-to-1 then looking at the inverse would give the opposite inequality:

$$\begin{aligned} \int_G f \circ \Phi \, J\Phi \, d\lambda &\leq \int_{\Phi(G)} f \circ \cancel{\Phi \circ \Phi^{-1}} \, \cancel{J\Phi \cdot J\Phi^{-1}} \, d\lambda \\ \int_{\Phi(G)} f \, d\lambda &\geq \int_G f \circ \Phi J\Phi \, d\lambda. \end{aligned}$$

Let $Q(c,r)$ be the cube $\prod_{i=1}^n [c_i - r_i, c_i + r_i]$. Consider what Φ does to this cube. Let $\Gamma^{(\delta)}$ be the set of points at most distance δ around Γ. By Taylor's Theorem with remainder,

$$\Phi(x) = \Phi(c) + \frac{\partial \Phi}{\partial x}(c)(x-c) + O(|x-c|^2).$$

There exist constant L and M so that

$$\Phi(Q(c,r)) \subseteq \left(\Phi(c) + T_{\frac{\partial\Phi}{\partial c}(x)}Q(c,r)\right)^{(Lr^2)}$$
$$\left(T_{\frac{\partial\Phi}{\partial x}(c)}Q(0,r)\right)^{(Lr^2)} \subseteq T_{\frac{\partial\Phi}{\partial x}(c)}Q(0,(1+Mr)r).$$

Thus taking Lebesgue measure

$$\lambda_{\mathbb{R}^N}(\Phi(Q(c,r))) \le (1+Mr)^N J\Phi(c)\lambda_{\mathbb{R}^N}(Q(c,r)).$$

Covering an open set with countably many nonoverlapping cubes Q with radius at most δ,

$$\lambda_{\mathbb{R}^N}(\Phi(H)) \le \sum_{Q\in C_\delta}(1+M\delta)^N J\Phi(c_Q)\lambda_{\mathbb{R}}(Q).$$

Thus $\lambda(\Phi(H)) \le \int_H J\Phi\, d\lambda$, i.e. (11) holds for f the characteristic function of an open set, hence for measurable sets. We have the result for characteristic, then simple, then measurable functions. □

Lecture 23
Fri. 4/1/11

§1 Polar coordinates

Let $\Phi : \mathbb{R}^N\backslash\{0\} \to \mathbb{S}^{N-1}$ be defined by $\Phi(x) = \frac{x}{|x|}$. Now

$$\lambda_{\mathbb{S}^{N-1}}(\Gamma) = N\lambda_{\mathbb{R}^N}(\{x : \Phi(x)\in\Gamma\}).$$

Theorem 23.1:

$$\int \varphi\, d\lambda_{\mathbb{R}} = \int_{(0,\infty)} \rho^{N-1}\left(\int_{\mathbb{S}^{N-1}} \varphi(\rho\omega)\,\lambda_{\mathbb{S}^{N-1}}(d\omega)\right) d\rho.$$

Proof. Let $f(r) = \int_{B(0,r)}\varphi\, dx$. Then

$$\begin{aligned}
f(r+h)-f(r) &= \int_{B(0,r+h)\backslash B(0,r)} \varphi\, dx\\
&= \int_{B(0,r+h)\backslash B(0,r)} \varphi(r\Phi(x))\,\lambda(dx) + \int_{B(0,r+h)\backslash B(0,r)} (\varphi(x)-\varphi(r\Phi(x)))\,\lambda(dx)\\
&= \int_{B(0,r+h)} \varphi(r\Phi(x))\, dx - \int_{B(0,r)} \varphi(r\Phi(x))\, dx + o(h)\\
&= (r+h)^N \int_{B(0,1)} \varphi(r\Phi(y))\, dy - r^N \int_{B(0,1)} \varphi(r\Phi(y))\, dy + o(h).
\end{aligned}$$

Therefore

$$f'(r) = \lim_{h \searrow 0} \frac{f(r+h) - f(r)}{h} = N r^{N-1} \int_{\mathbb{S}^{N-1}} \varphi(r\omega)\, \lambda_{\mathbb{S}^{N-1}}(d\omega).$$

Now integrate from 0 to ∞. □

The polar coordinate formula can be viewed as a special case of surface measure.

§2 Surface measure

Definition 23.2: A **hypersurface** is a set $M \subseteq \mathbb{R}^N$ such that for every $p \in M$ there exists $r > 0$, $F \in C^3(B(p,r);\mathbb{R})$, so that $|\nabla F| > 0$ and

$$M \cap B(p,r) = \{x \in B(p,r) : F(x) = 0\}.$$

The tangent space $T_p(M)$ is the set of $v \in \mathbb{R}^N$ such that there exists $\varepsilon > 0$ so there exists $\geq \in C^1((-\varepsilon, \varepsilon), M)$ such that so that $\gamma(0) = p$, $\dot{\gamma}(0) = v$.

In the language of manifolds, M is a submanifold of $\mathbb{R}^n$, locally characterized by the fact that the 0 is a regular point of F.

Lemma 23.3: 1. M is a countable union of compact sets.

2. $T_p(M)$ is a $N-1$ dimensional subspace of $\mathbb{R}^N$ and $T_p(M) = \{v \in \mathbb{R}^N : v \perp \nabla F(p)\}$.

Proof. For the second part, noting the directional derivative is given by dot product with the gradient,

$$0 = \frac{d}{dt} F(\gamma(t))|_{t=0} = \langle \nabla F(p), v \rangle.$$

Conversly, suppose $\langle \nabla F(p), v \rangle = 0$. By ODE theory (existence of solutions), we can find γ locally so that $\dot{\gamma}(0) = v$ and

$$\dot{\gamma}(x) = v - \frac{\langle v, \nabla F(x) \rangle}{|\nabla F(\gamma(t))|^2} \nabla F(x).$$

Now $\frac{d}{dt} F(\gamma(t)) = \langle \dot{\gamma}(t), \nabla F(\gamma(t)) \rangle = 0.$

□

Lecture 24
Mon. 4/4/11

§1 Surface measure

Definition 24.1: Define a cylindrical neighborhood of a hypersurface M to be

$$\Gamma(\rho) = \{x : \exists p \in \Gamma, |x - \rho| < \rho,\ x - p \perp T_p(M).\}$$

We will define a measure by

$$\lambda_M(\Gamma) = \lim_{\rho \searrow 0} \frac{\lambda_{\mathbb{R}^N}(\Gamma(\rho))}{2\rho}.$$

Now

$$\Gamma(\rho) = \left\{ q + \xi \frac{\nabla F(q)}{|\nabla F(q)|} : q \in \Gamma,\ |\xi| < \rho \right\}.$$

This is measurable because a Lipschitz function maps Lebesgue measurable sets to measurable sets (it maps compact to compact, closed to closed, countable unions of closed to closed; every Lebesgue measurable set differs by at most a set of measure 0 from such a set; by Lipschitz continuity a set of measure 0 goes to a set of measure 0), and $(q, \xi) \mapsto q + \xi \frac{\nabla F(q)}{|\nabla F(q)|}$ is Lipschitz.

Proposition 24.2: For every $p \in M$ there exists U an open neighborhood of $O_{\mathbb{R}^{N-1}}$ and a $\Psi : U \to M$ such that $\Psi(O_{\mathbb{R}^{N-1}}) = p$ and there exists $r > 0$ such that $\Psi(U) = M \cap B_{\mathbb{R}^N}(p, r)$, and for every $u \in U$

$$\{\partial_{u_1}\Psi(u), \ldots, \partial_{u_{N-1}}\Psi(u)\}$$

is a basis for $T_{\Psi(u)}(M)$. ((U, Ψ) is a coordinate chart.)

Proof. By taking r small enough we may assume without loss of generality that $|\partial_{x_N} F| > 0$ on $B_{\mathbb{R}^N}(p, r)$. Define

$$\Phi(x) = \begin{bmatrix} x_1 - p_1 \\ \cdots \\ x_{N-1} - p_{N-1} \\ F(x) \end{bmatrix}.$$

The Jacobian is

$$\frac{\partial \Phi}{\partial x} = \begin{pmatrix} I_{N-1} & & 0 \\ \partial_{u_1} F & \cdots & \partial_{u_N} F \end{pmatrix}.$$

Its determinant doesn't vanish so by the Inverse Function Theorem, F admits an inverse in a neighborhood of the origin. Define $\Psi(u) = \Phi^{-1}((u, 0))$ on a neighborhood of 0. Note $\Phi^{-1}(O_{\mathbb{R}^{N-1}}) = p$, and $\Psi(u) \in M$. Thus the derivative is in the tangent space; each $\partial_{u_i}\Psi(u)$ is in the tangent space. Since they are linearly independent and there are $N - 1$ of them, they form a basis. (The $N - 1$ column vectors are the first $N - 1$ vectors in the Jacobian of Φ^{-1}, which has nonzero determinant.) $\square$

Let (U, Ψ) be a coordinate chart at p. Define $\tilde{\Psi}$ on $U \times \mathbb{R}$ by

$$\tilde{\Psi}(u, \xi) = \Psi(u) + \xi n(u)$$

where $n(u)$ is the normal at u. It parameterize points near the hypersurface by points on the hypersurface and how far to go perperdicularly. Note

$$(\det A)^2 = \det(A^T A) = \det\left[(v_i, v_j)_{\mathbb{R}^N}\right]_{1 \le i,j \le N}.$$

The Jacobian of $\tilde{\Psi}$ at $(u, 0)$ is

$$J\tilde{\Psi}(u,0) = \sqrt{\left[\langle \partial_{u_i}\Psi, \partial_{u_j}\Psi\rangle\right]_{1\le i,j\le N}} > 0$$

since ???

$$J\tilde{\Psi}(u,0) = (\partial_{u_1}\Psi, \ldots, \partial_{u_{N-1}}\Psi, n\Psi(u)).$$

Thus $\tilde{\Psi}$ is locally invertible. FIXFIXFIXFIXFIXFIXFIXFIXFIXFIXFIXFIXFIXFIXFIX-FIXFIXFIX

For $\Gamma \subseteq M \cap B(p, r)$, by changes of variables,

$$\begin{aligned}
\lambda(\Gamma(p)) &= \int_{\tilde{\Psi}^{-1}(\Gamma(\rho))} 1_{\Gamma(\rho)}(\tilde{\Psi}(u,\xi))\, J\tilde{\Psi}(u,\xi). \\
&= \int_{\tilde{\Psi}^{-1}(\Gamma)\times(-\rho,\rho)} J\tilde{\Psi}(u,\xi)\, d\xi du \\
&= \int_{\Psi^{-1}(\Gamma)} \left(J\tilde{\Psi}(u,\xi)\, d\xi \right)\, du.
\end{aligned}$$

Divide by 2ρ and take the limit as $\rho \to 0$ to get

$$\int_{\Phi^{-1}(\Gamma)} J\tilde{\Psi}(u,0)\, du.$$

The limit exists and is given by a measure!

In differential geometry we go in the opposite direction: define a hypersurface (or manifold) to be something which has a coordinate chart at each point. Then define a measure on the hypersurface by using the formula above. We don't usually talk about the relationship to the measure in the ambient space because we use intrinsic descriptions.

If $\mathbb{R}^N \to \mathbb{R}^N$ is Lipschitz continuous $|V(y) - V(x)| \le L|y - x|$ then we can always solve $\frac{d\Phi}{dt}(t,x) = V(\Phi(t,x))$ and $\Phi(0,x) = x$.

Flow property: $\Phi(s+t,x) = \Phi(t,\Phi(s,x))$. (Use uniqueness of solutions.)

Lecture 25
Wed. 4/6/11

§1 Divergence

Definition 25.1: The **divergence** of $V : \mathbb{R}^N \to \mathbb{R}^N$ is

$$\operatorname{div}(V) = \sum_{i=1}^{N} \partial_{x_i} V_i.$$

Theorem 25.2 (Divergence theorem)**:** Let G be a smooth bounded domain in $\mathbb{R}^N$, i.e. it is an open set and its boundary is a hypersurface. Let $V : \mathbb{R}^N \to \mathbb{R}^N$ be a function with two bounded continuous derivatives. Then

$$\int_G \operatorname{div}(V)\, dx = \int_{\partial G} \langle V, n\rangle \, d\lambda_M$$

where n is the outward pointing normal.

Proof. Let Φ be the flow of V, i.e. $\dot{\Phi}(t,x) = V(\Phi(t,x))$, $\Phi(0,x) = x$. The flow property states that $\Phi(s+t,x) = \Phi(t,\Phi(s,x))$. Thus $I = \Phi(t,\Phi(-t,x))$; the $\Phi(t,\cdot)$'s are bijective, in fact diffeomorphisms. Differentiating $\dot{\Phi}(t,x) = V(\Phi(t,x))$ with respect to x,

$$\partial_t \frac{\partial \Phi}{\partial x}(t,x) = \frac{\partial V}{\partial y}\Phi(t,x)\frac{\partial \Phi}{\partial x}(t,x). \tag{12}$$

We are interested in $\partial_t \det(\Phi)$.

For a matrix $A = ((a_{ij}))$, $\partial_{a_{ij}} \det(A) = A^{(ij)}$ where $A^{(ij)}$ is the (i,j)th cofactor of A. From (12), and $\sum_{j=1}^N a_{kj}A^{(ij)} = \det(A)$, Using the chain rule,

$$\begin{aligned}
\frac{d}{dt}\det\left(\frac{\partial \Phi(t,x)}{\partial x}\right) &= \sum_{1\le i,j,k\le N} \left(\frac{\partial \Phi(t,x)}{\partial x}\right)^{(ij)} \left(\frac{\partial V}{\partial y}(\Phi(t,x))\right)_{ik} \left(\frac{\partial \Phi(t,x)}{\partial x}\right)_{kj} \\
&= \operatorname{div}(V)(\Phi(t,x))\det\left(\frac{\partial \Phi(t,x)}{\partial x}\right).
\end{aligned}$$

Solving this differential equation, and using $\frac{1}{J\Phi(t,x)} = J\Phi(-t,\Phi(t,x))$ gives

$$\begin{aligned}
\det\left(\frac{\partial \Phi}{\partial x}\right) &= e^{\int_0^t \operatorname{div}(V)(\Phi(t,x))\,dt} \\
\int \varphi\circ\Phi(t,x)\,dx &= \int \varphi(y)J\Phi(-t,y)\,dy \\
&= \int \varphi(y)e^{-\int_0^t \operatorname{div}(V)(\Phi(-t,y))}\,dt\,dy.
\end{aligned}$$

Now consider

$$\lambda_{\mathbb{R}}(G) - \lambda_{\mathbb{R}^N}(\Phi(t,G)) = \int_G (1 - e^{\int_0^{-t} V(\Phi(t,x))\,dt})dx.$$

So

$$\lim_{t\to 0}\frac{\lambda_{\mathbb{R}}(G) - \lambda_{\mathbb{R}^N}(\Phi(H,G))}{t} = ?$$

Now we calculate another way.

$$\begin{aligned}
\lambda_{\mathbb{R}}(G) - \lambda_{\mathbb{R}^N}(\Phi(t,G)) &= \int 1_{G(x)} - 1_G(\Phi(t,x))\,dx \\
&= \int_G 1_{G^c}(\Phi(t,x))\,dx - \int_{G^c} 1_G(\Phi(t,x))\,dx.
\end{aligned}$$

Do something. Then

$$\begin{aligned}\langle n\Phi(t,x),\Phi(t,x)-p\Phi(t,x)\rangle &= \langle n(x),\Phi(t,x)-p\Phi(t,x)\rangle + O(t^2)\\ &= \langle n(x),\Phi(t,x)-x\rangle + \langle n(x),x-p(x)\rangle\\ &\quad + \langle n(x),p(x)-p(\Phi(t,x))\rangle + O(t^2)\\ &= t\langle n(x),V(x)\rangle + \langle n(x),x-p(x)\rangle\\ &\quad + \langle n(x),p(x)-p(\Phi(t,x))\rangle + O(t^2)\\ &= t\langle n(x),V(x)\rangle + \langle n(x),x-p(x)\rangle + E(t,x).\end{aligned}$$

Note we used $\langle n(x), p(\Phi(t,x))-p(x)\rangle$ is 0 when $t=0$ and the derivative with respect to t is 0 since the second argument is on the boundary, so its derivative is $t=0$ in the tangent space, perpendicular to $n(x)$, so is $O(t^2)$. □

Lecture 26
Fri. 4/8/11

§1 Jensen's inequality

Definition 26.1: A subset $C\subseteq\mathbb{R}^N$ is **convex** if for all $x,y\in C$ and $0\le t\le 1$, $(1-t)x+ty\in C$. I.e. if it contains two points it contains the line segment joining them.

A function $g: C\to\mathbb{R}$ is **concave** if $g((1-t)x+ty)\ge(1-t)g(x)+tg(y)$.

Proposition 26.2: g is concave iff $\{(x,\xi): x\in C \text{ and } \xi\le g(x)\}$ is convex in $\mathbb{R}^N\times\mathbb{R}$.

Theorem 26.3 (Jensen's inequality, discrete version)**:** Let α_i be nonnegative numbers summing to 1. Then

$$g\left(\sum_{k=1}^n \alpha_k x_k\right)\ge\sum_{k=1}^n\alpha_k g(x_k).$$

Proof. $k=2$ case is definition. Induct. □

We think of this as a statement about probability measures: Think of the LHS as $\int_C x\,\mu(dx)$, where $\mu=\sum_{k=1}^N\alpha_k(x)\delta_{x_k}(x)$, δ_x being the measure assigning 1 to $\{x\}$ and 0 to $\mathbb{R}^N-\{x\}$.

Theorem 26.4 (Jensen's inequality, continuous version)**:** Let $(E,\mathcal{B},\mu)$ be a probability space (i.e. a measure space with $\mu(E)=1$), C be a closed, convex set in $\mathbb{R}^N$, and $g: C\to[0,\infty)$ be continuous and concave. If $F: E\to C$ is $\mathcal{B}$-measurable such that $|F|\in L^1(\mu)$ then $\int F\,d\mu\in C$ and

$$g\left(\int F\,d\mu\right)\ge\int g\circ F\,d\mu.$$

Proof. First assume F only takes a finite number of values, $y_0, \ldots, y_n$. Then $\int F\,d\mu = \sum_{k=0}^{n} y_k \mu(F = y_k) \in C$, and the inequality holds by discrete Jensen's.

Now we approximate F by simple functions. First we get rid of large values of F. Given $m \geq 1$, choose $R_m < \infty$ such that $\int_{|F|\geq R_m} |F|\,d\mu \leq \frac{1}{m}$. (Possible since $|F|$ is integrable.) The set $C \cap \overline{B(0,R)}$ is compact so it can be covered with a finite number of balls of radius $\frac{1}{m}$, say as $\bigcup_{k=1}^{n_m} B(y_k, \frac{1}{m})$.

Choose any $y_0 \in C$. Define

$$F_m(e) = \begin{cases} y_k & \text{if } F(e) \in B\left(y_k, \frac{1}{m}\right) \setminus \bigcup_{j=0}^{k-1} B\left(y_j, \frac{1}{m}\right) \\ y_0 & \text{if } |F(e)| > R_m. \end{cases}$$

Since F_m is simple, $\int F_m\,d\mu \in C$, and $g\left(\int F_m\,d\mu\right) \geq \int g \circ F_m\,d\mu$. Since $\int |F - F_m|\,d\mu \leq \frac{2}{m}$ (the integral is at most $\frac{1}{m}$ on the set where $F(x) \leq R_m$ and where $F(x) > R_m$), $\lim_{m\to\infty} \int |F - F_m|\,d\mu = 0$. Hence $\int F_m\,d\mu \to \int F\,d\mu$. Since C is closed, $\int F\,d\mu \in C$. Consider

$$g\left(\int F_m\,d\mu\right) \geq \int g \circ F_m\,d\mu.$$

The LHS converges to $g\left(\int F\,d\mu\right)$ by continuity of g and $\int F_m\,d\mu \to \int F\,d\mu$, and the RHS converges to $\int g \circ F\,d\mu$ by Fatou's lemma. □

Theorem 26.5: Let U be open and convex. Then $C = \overline{U}$ is also convex. Suppose $g : C \to \mathbb{R}$ is C^2 on U. Then g is concave on C if and only if

$$H_g(x) = \left[\frac{\partial^2 g}{\partial x_i \partial y_j}(x)\right]_{1\leq i,j\leq N}.$$

has all eigenvalues nonpositive, for all $x \in U_i$.

Proof. Suppose $H_g(x)$ has a positive eigenvalue at $e \in \mathbb{S}^{N-1}$. Then there is e so that $\langle e, H_g(x)e\rangle > 0$. Let $u(t) = g(x + te)$. Then

$$\frac{d^2}{dt^2} g(x + te)|_{t=0} = \langle e, H_g(x)e\rangle > 0.$$

This implies $g(x + te) + g(x - te) > g(x)$ for small x, so g is not concave.

Conversely, note that if u is a C^2 function such that $u(0) = u(1) = 0$ and $u''(t) \leq 0$, then $u(t) \geq 0$ on $[0, 1]$. Apply this to $u(t) = g((1 - t)x + ty) - (1 - t)g(x) - tg(y)$ where $y - x$ ranges over all directions. □

To apply this in the two dimensional case, note that $A = \left[\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right]$ has all eigenvalues nonpositive iff $\mathrm{Tr}(A) \leq 0$ and $\det(A) \leq 0$.

Suppose $\alpha \in (0, 1)$. Consider $x^\alpha y^{1-\alpha}$ on $x, y \geq 0$. This is concave on $[0, \infty)^2$. So is $(x^\alpha + y^\alpha)^{\frac{1}{\alpha}}$.

Theorem 26.6 (Minkowski's inequality)**:** Let $(E, \mathcal{B}, \mu)$ be a probability space and $f, g : E \to [0, \infty)$ be $\mathcal{B}$-measurable. Let $p \in (1, \infty)$. Then

$$\left(\int (f+g)^p \, d\mu\right)^{\frac{1}{p}} \le \left(\int f^p \, d\mu\right)^{\frac{1}{p}} + \left(\int g^p \, d\mu\right)^{\frac{1}{p}}.$$

Proof. Apply Jensen to $(x^\alpha + y^\alpha)^{\frac{1}{\alpha}}$. □

Theorem 26.7 (Hölder's inequality)**:** Let $p \in (1, \infty)$, $p' = \frac{p}{p-1}$ (note $\frac{1}{p} + \frac{1}{p'} = 1$),

$$\int fg \le \left(\int f^p\right)^{\frac{1}{p}} \left(\int g^{p'}\right)^{\frac{1}{p'}}$$

Considering $\frac{\mu}{\mu(E)}$ gives the above inequalities for finite measure spaces. We can extend to σ-finite measure spaces too.

We will introduce a norm on $L^p(\mu; \mathbb{R})$, $\|f\|_{L^p} = \left(\int |f|^p\right)^{\frac{1}{p}}$. Minkowski's inequality gives that this is a valid norm (triangle inequality holds). Again we look at equivalence classes.

Lecture 27
Mon. 4/11/11

§1 Lebesgue spaces

For $p \in [1, \infty)$, let $L^p(\mu, \mathbb{R})$ be the set of measurable functions $\xi : E \to \mathbb{R}$ such that

$$\|\xi\| = \left(\int |f|^p\right)^{\frac{1}{p}} < \infty.$$

Note that

$$\begin{aligned} \|\xi + \psi\|_p &\le \|\xi\|_p + \|\psi\|_p \\ \left| \|\xi\|_p - \|\psi\|_p \right| &\le \|\xi - \psi\|_p \\ \|\alpha\xi\|_p &= |\alpha| \, \|\xi\|_p \end{aligned}$$

Again we form equivalence classes, considering two functions the same if they differ on a set of measure 0. This turns $L^p(\mu, \mathbb{R})$ into a metric space; it satisfies the triangle inequality by Minkowski's inequality (??).

Let $L^\infty(\mu, \mathbb{R})$ be the set of measurable functions that are bounded Off a set of measure 0. Define the norm to be

$$\|f\|_\infty = \inf\{M : |f| \le M(\text{ a.e., } \mu)\}.$$

Some intuition.

1. As p changes, we emphasize different features of the function. As p gets large, large values f takes ("spikes") are emphasized. Small p is forgiving on spikes.

2. If the space is infinite, a function in L^p for small p must have rapid "decay". If p is large and $|f| < 1$, $|f|^p$ is small, so f need not decay as rapidly. For example $\frac{1}{1+|x|}$.

3. Often you can't get pointwise estimates of functions (ex. solutions to differential equations). However, using various trick, you can get estimates of integrals of the functions. We want the flexibility of being able to choose p. The Sobolev imbedding theorem gives an inequality for $\|f\|_p$ given that $f \in L^p(\mathbb{R}^N, \mathbb{R})$, $|\nabla f| \in L^p$, and $p > N$.

4. Some spaces are geometrically nicer than others. For example, for L^1, for a space on two points, a circle is a diamond and has nasty corners. For $1 < p < \infty$ it is smooth and convex. For $p = \infty$ we again get nastiness (a square).

Theorem 27.1: 1. If $f_n \to f$ in L^p, then $f_n \to f$ in μ-measure.

2. If $f_n \to f$ in μ-measure of μ-almost everywhere. Then
$$\|f\|_p \leq \varliminf_{n\to\infty} \|f_n\|_p \,.$$

3. $\sum_{n>m} \|f_n - f_m\|_p \to 0$ as $m \to \infty$, then f_n converges to a function in L^p. Hence L^p is a complete metric space.

4. If μ is finite, $\mathcal{B} = \sigma(\mathcal{C})$ is a Π-system, and the α_m are dense in $\mathbb{R}$, then
$$S = \left\{ \sum_{m=1}^{n} \alpha_m 1_{\Gamma_m} \right\}$$
is dense in L^p for $p < \infty$. If the measure space is σ-finite and the σ-algebra can be generated by a countable set, then L^p is separable for $p < \infty$.

5. Suppose (E, ρ) is a metric space and $\{E_n : n \geq 1\}$ is a countable set such that $E_n \nearrow E$, with $\mu(E_n) < \infty$, $p < \infty$. Then the set of continuous functions which vanish off E_n for some n is densein L^p.

6. (Lebesgue dominated convergence) Suppose $p < \infty$. If $f_n \to f$ in μ-measure or μ-almost everywhere and there exists $g \in L^p$ such that $|f_n| \geq g$ almost everywhere for every n. Then $f_n \to f$ in L^p.

7. (Lieb) if $f_n \to f$ in μ-measure or μ-almost everywhere, and $\sup_n \|f_n\|_1 < \infty$, then
$$\lim_{n\to\infty} \int ||f_n|^p - |f|^p - |f - f_n|^p| \; d\mu = 0$$
Thus if $\|f_n\|_p \to \|f\|_p$, then $\|f_n - f\|_p \to 0$.

Proof. The arguments are the same as for L^1.

For Lieb, note there is K_p such that

$$||b|^p - |a|^p - |b-a|^p| \le K_p(|a||b-a|^{p-1} + |a|^{p-1}|b-a|)$$

for all a, b. (This follows from homogeneity and $||c|^p - 1 - |c-1|^p| \le K_p(|c-1|^{p-1} + |c-1|)$. Use calculus!) Put $a = f_n(x)$ and $b = f(x)$. Now integrate and note that if $\xi_n \in L^p$, $\xi_n \to 0$ in μ-measure or a.e., then

$$\lim_{n\to\infty} \int |\xi_n|^{p-1}|\psi|\, d\mu = 0 = \lim_{n\to\infty} \int |\xi_n||\psi|^{p-1}\, d\mu.$$

To prove this, assume they're positive. Now divide into 2 parts. For $\delta > 0$,

$$\begin{aligned}\int \xi_n^{p-1}\psi &= \int_{\xi_n \le \delta\psi} \xi_n^{p-1}\psi\, d\mu + \int_{\xi_n > \delta\psi} \xi_n^{p-1}\psi\, d\mu \\ &\le \delta^{p-1}\, \|\psi\|_p^p + \int_{\varphi_m \ge \delta^2} \xi_n^{p-1}\psi\, d\mu + \int_{\psi \le \delta} \xi_n^{p-1}\psi\, d\mu\end{aligned}$$

Take $p' = \frac{p}{p-1}$ and use Hölder's inequality:

$$\int_{\varphi_m \ge \delta^2} \xi_n^{p-1}\psi\, d\mu + \int_{\psi \le \delta} \xi_n^{p-1}\psi\, d\mu \le \left(\int \xi_n^p\right)^{\frac{p-1}{p}} \left[\left(\int_{\xi_n \ge \delta^2} |\psi|^p\right)^{\frac{1}{p}} + \left(\int_{\psi \le \delta} \psi^p\, d\mu\right)^{\frac{1}{p}}\right]$$

Now let $n \to \infty$, $\delta \to 0$, and use Lebesgue's Dominated Convergence Theorem. The other inequality follows similarly. □

Lecture 28
Wed. 4/13/11

§1 Inequalities, continued

Theorem 28.1: For p and p' in $(1, \infty)$ such that $\frac{1}{p} + \frac{1}{p'} = 1$,

$$\|fg\|_1 \le \|f\|_p \|g\|_{p'}.$$

For $f \in L^p$,

$$\|f\|_p = \sup\{\|fg\|_1 : g \in L^{p'},\ \|g\|_{p'} \le 1\}$$

Proof. The first equality is just Hölder's inequality. For the second statement, "$\le$" follows from the first equality. If $\|f\|_p = 0$ then f vanishes except on a set of measure 0 so equality holds. Else, take $g = \frac{|f|^{p-1}}{\|f\|_p^{p-1}}$. Then $|g|^{p'} = \frac{|f|^p}{\|f\|^p}$ and $|fg| = \frac{|f|^p}{\|f\|_p^{p-1}}$. (For $p = 1$ take $g = 1$.) □

The following gives a similar result when it is not *a priori* known that $f \in L^p$.

Theorem 28.2: If $p = 1$; or $p \in (1, \infty)$ and either $\mu(|f| > \delta) < \infty$ or for every $\delta > 0$ or μ is σ-finite, then

$$\|f\|_p = \sup\{\|fg\|_1 : g \in L^{p'},\ \|g\|_{p'} \le 1\}.$$

§2 Mixed Lebesgue spaces

Definition 28.3: let $(E_1, \mathcal{B}_1, \mu_1)$ and $(E_2, \mathcal{B}_2, \mu_2)$ be σ-finite. Suppose f is $\mathcal{B}_1 \times \mathcal{B}_2$-measurable, $p_1, p_2 \in [1, \infty)$. Then define

$$\|f\|_{L^{(p_1,p_2)}(\mu_1 \times \mu_2, R)} = \left\| \|f(x_1, x_2)\|_{L^{p_1}(\mu_1;\mathbb{R})} \right\|_{L^{p_2}(\mu_2;\mathbb{R})}$$
$$= \left[\int \left(\int |f(x_1, x_2)|^{p_1} \, \mu(dx_1) \right)^{\frac{p_2}{p_1}} \, \mu(dx_2) \right]^{\frac{1}{p_2}}.$$

($L^{(p_1,p_2)}$ denotes the space of functions for which the above norm is defined and finition.) Note for $p_1 = p_2$ this defined just like in Fubini.

Theorem 28.4 (Generalized Minkowski's inequality)**:** If $p_1 \le p_2$, then

$$\|f\|_{(p_1,p_2)} \le \|f\|_{(p_2,p_1)}.$$

For example, taking $E_1 = \{1, 2\}$, $p_1 = 1$ and $p_2 = p$ we get out Minkowski's inequality. This will give estimates for linear operators on L^p spaces. First we prove need the following lemma.

Lemma 28.5: For μ_1 and μ_2 finite, for every $\varepsilon > 0$ there exists a function $\psi(x_1, x_2) = \sum_{m=1}^n 1_{\Gamma_{1,m}}(x_1)\varphi_m(x_2)$ where

1. the $\Gamma_{1,m}$'s are mutually disjoint,
2. the φ_m are bounded, and
3. $\|f - \psi\|_{(p_1,p_2)} < \varepsilon$.

We want to approximate f on $E_1 \times E_2$ by these functions because the fact that the $\Gamma_{1,m}$ are mutually disjoint makes it easy to raise ψ to a power:

$$\left(\sum_{m=1}^n 1_{\Gamma_{1,m}}(x_1)\varphi_m(x_2) \right)^p = \sum_{m=1}^n 1_{\Gamma_{1,m}}(x_1)\varphi_m(x_2)^p$$

Proof. By Hölder's inequality with $\frac{q}{p}$, $p \le q$,

$$\int_E |f|^p \, d\mu \le \left(\int_E |f|^{p \cdot \frac{q}{p}} \, d\mu \right)^{\frac{p}{q}} \left(\int_E 1 \, d\mu \right)^{1-\frac{p}{q}} = \left(\int_E |f|^q \, d\mu \right)^{\frac{p}{q}} \mu(E)^{1-\frac{p}{q}}.$$

Thus there exists C so that

$$\|g\|_{(p_1,p_2)} \le C \, \|g\|_{L^q(\mu_1 \times \mu_2)}, \tag{13}$$

and convergence of functions in $L^q(\mu_1\times\mu_2)$ implies convergence in $L^{(p_1,p_2)}$. We know functions of the form $\sum_{k=1}^l a_j 1_{\Gamma_{1,j}\times\Gamma_{2,j}}$ are dense in $\mu_1\times\mu_2$ since sets of the form $\Gamma_1\times\Gamma_2$ generate the σ-algebra (Theorem 27.1(4)). Let $\eta\in I:=\{0,1\}^l$ and let

$$\Gamma_{1,\eta}=\Gamma_{1,1}^{(\eta_1)}\cap\cdots\cap\Gamma_{1,l}^{(\eta_2)}.$$

(Define $\Gamma^{(1)}=\Gamma$, $\Gamma^{(0)}=\Gamma^c$.) Then

$$\begin{aligned}
1_{\Gamma_{1,j}}(x_1) &= \sum_{\eta\in I}\eta_j 1_{\Gamma_{1,\eta}}(x_1)\\
\sum_{j=1}^l a_j 1_{\Gamma_{1,j}\times\Gamma_{2,j}} &= \sum_{j=1}^l\sum_{\eta} a_j\eta_j 1_{\Gamma_{1,\eta}}(x_1)1_{\Gamma_{2,\eta}}(x_2)\\
&= \sum_{\eta\in I} 1_{\Gamma_{1,\eta}}(x_1)\underbrace{\sum_{j=1}^l \eta_j a_j 1_{\Gamma_{2,j}}(x_2)}_{\varphi_\eta}
\end{aligned}$$

□

Proof. (of Theorem 28.4) It suffices to show the finite case since a σ-finite measure is the union of subsets of finite measure. By Lemma 28.5 and (13), it suffices to show the result for functions in the above form.

Now (using Minkowski, noting $p_2\geq p_1$)

$$\begin{aligned}
\|\psi\|_{(p_1,p_2)} &= \left[\int_{E_2}\left(\sum_{m=1}^n \mu_1(\Gamma_{1,m})|\varphi_m(x_2)|^{p_1}\,\mu(dx_1)\right)^{\frac{p_2}{p_1}}\mu(dx_2)\right]^{\frac{1}{p_2}}.\\
&\leq \left[\sum_{m=1}^n \mu_1(\Gamma_{1,m})\,\|\varphi_m\|_{L^{p_2}}^{p_1}\,\mu_2(dx_2)\right]^{\frac{1}{p_1}}\\
&= \|\psi\|_{(p_2,p_1)}
\end{aligned}$$

□

We investigate linear maps $\mathcal{K}:L^p(\mu)\to L^q(\nu)$. We think of $K:E_2\times E_1\to\mathbb{R}$ as a "matrix representation" of $\mathcal{K}$ if

$$\mathcal{K}f(x_2)=\int K(x_2,x_1)f(x_1)\,\mu_1(dx_1).$$

We estimate this in terms of the function K (kernel).

Lecture 29
Fri. 4/15/11

§1 Transformations on Lebesgue Spaces

Proposition 29.1: If f and g are functions on $E_1 \times E_2$ and $\frac{1}{p} + \frac{1}{p'} = \frac{1}{q} + \frac{1}{q'} = 1$, then

$$\|fg\|_{L^1} \le \|f\|_{(p,q)} \|g\|_{(p',q')}.$$

Proof. Two applications of Hölder's inequality. □

We think of $K : E_2 \times E_1 \to \mathbb{R}$ as a "matrix representation" of $\mathcal{K}$ if Consider the map $\mathcal{K} : L(\mu_2) \to L(\mu_1)$ defined by

$$(\mathcal{K}f)(x_1) = \int K(x_1, x_2) f(x_2)\, \mu_2(dx_2).$$

Think of K as a matrix representation of $\mathcal{K}$. K is called the kernel.

Any linear map $f : \mathbb{R}^m \to \mathbb{R}^n$ is continuous: Write $x = \sum_{j=1}^m \langle x, e_j \rangle e_j$; then $f(x) = \sum_{j=1}^n \langle x, e_j \rangle f(e_j)$. However this is not necessarily true for infinite-dimensional spaces.

To check a linear functions it suffices to check continuity at 0. Thus to check that $\mathcal{K} : L^p(\mu_2) \to L^p(\mu_1)$ is continuous, if suffices to show that

$$\|\mathcal{K}\|_{L^r(\mu_1)} \le C \|f\|_{L^p(\mu_2)}. \tag{14}$$

Theorem 29.2: Let $(E_1, \mu_1), (E_2, \mu_2)$ be σ-finite measure spaces, and let $p \in [1, \infty]$ and $q \in [1, \infty)$. Let r be so that $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1 \ge 0$. Define

$$\begin{aligned} M_1 &= \sup_{x_2 \in E_2} \|K(x_1, x_2)\|_{L^q(\mu_1)} < \infty \\ M_2 &= \sup_{x_1 \in E_2} \|K(x_1, x_2)\|_{L^q(\mu_2)} < \infty. \end{aligned}$$

Given $f \in L^{q'}(\mu_2)$, define

$$(\mathcal{K}f)(x_1) = \int K(x_1, x_2) f(x_2)\, \mu_2(dx_2).$$

(The integral is finite by Hölder.) Then

$$\|\mathcal{K}f\|_{L^r(\mu_1)} \le M_1^{\frac{q}{r}} M_2^{1-\frac{q}{r}} \|f\|_{L^p(\mu_2)}.$$

Proof. First suppose $r = \infty$. Then $p = q'$. Therefore by Hölder's inequality,

$$\begin{aligned} \left| \int K(x_1, x_2) f(x_2)\, \mu_2 \right| &\le \|K(x_1, x_2)\|_{L^q(\mu_2)} \|f\|_{L^{q'}(\mu_2)} \\ &\le M_2 \|f\|_{L^p(\mu_2)} \end{aligned}$$

Now take the sup.

Now suppose $p = 1$. Then $q = r$. Then by Generalized Minkowski (28.4),

$$\begin{aligned}
\|\mathcal{K}f\|_{L^r(\mu_1)} &\leq \|Kf\|_{L^{(1,r)}(\mu_2,\mu_1)} \\
&\leq \|Kf\|_{L^{(r,1)}(\mu_1,\mu_2)} \\
&= \int \left(\int |K(x_1,x_2)|^r |f(x_2)|^r \, \mu_1(dx_1) \right)^{\frac{1}{r}} \mu_2(dx_2) \\
&= M_1 \|f\|_{L^1(\mu_2)} .
\end{aligned}$$

Switching order of integration with Minkowski allows $f(x_2)$ to come out as a constant.

Now suppose $r < \infty$, $p > 1$. Then $q < r$ and $p \leq r$. Let $\alpha = \frac{q}{r} \in (0,1)$. Then $(1-\alpha)p' = q$. From (28.1), since $(E_1, \mu_1), (E_2 \mu_2)$ are σ-finite,

$$\|\xi\|_{L^r} = \sup\{\|\xi\psi\| : \psi \in L^{r'}, \|\psi\|_{r'} \leq 1\}.$$

For $g \in L^{r'}(\mu_1)$,

$$\begin{aligned}
\|g\mathcal{K}f\|_{L^1(\mu_1)} &\leq \|\mathcal{K}\|_{L^1(\mu_1 \times \mu_2)} \\
&\leq \left\| (|f||K|^\alpha)(|K|^{1-\alpha}|g|) \right\|_1 \\
&\leq \||K|^\alpha f\|_{(r,p)} \left\| g|K|^{1-\alpha} f \right\|_{(r',p')} \\
&\leq \||K|^\alpha f\|_{(r,p)} \left\| g|K|^{1-\alpha} f \right\|_{L^{(p',r')}(\mu_2,\mu_1)} && \text{by Minkowski}
\end{aligned}$$

Now

$$\begin{aligned}
\||K|^\alpha f\|_{(r,p)} &= \left[\int \left(\int |f(x_2)|^r |K(x_1,x_2)|^q \, \mu(dx_1) \right)^{\frac{p}{r}} \mu_2(dx_2) \right]^{\frac{1}{p}} . \\
&\leq M_1^{\frac{q}{r}} \|f\|_{L^p(\mu_2)} .
\end{aligned}$$

Similarly,

$$\left\| g|K|^{1-\alpha} \right\|_{(p',r')} \leq M_2^{1-\alpha} \|g\|_{L^{r'}(\mu_2)} .$$

Combining these two gives the desired result. □

Note in the case of matrices, letting $A = [a_{i,j}]$,

$$\begin{aligned}
M_1 &= \max_j \left(\sum_{i=1}^n |a_{ij}|^q \right)^{\frac{1}{q}} \\
M_2 &= \max_j \left(\sum_{j=1}^n |a_{ij}|^q \right)^{\frac{1}{q}}
\end{aligned}$$

Then

$$M_1^{\frac{q}{r}} M_2^{1-\frac{q}{r}} \left(\sum_j |x_j|^p \right)^{\frac{1}{p}}.$$

Theorem 29.2 gives (14), hence continuity of $\mathcal{K}$.

Theorem 29.3: $\mathcal{K}$ is continuous.

Lecture 30
Wed. 4/20/11

§1 Convolution

Theorem 29.2 gives the following.

Theorem 30.1: Define

$$\Lambda_k(f) = \left\{ x : \int |K(x_1, x_2)||f(x_2)| \, \mu_2(dx_2) < \infty \right\}.$$

Then

$$\mu_1(\Lambda_k(f)^c) = 0.$$

In addition, defining

$$(\overline{\mathcal{K}}f)(x_1) \begin{cases} \int K(x_1, x_2) f(x_2) \, \mu_2(dx_2), & x_1 \in \Lambda_k(f) \\ 0, & x_1 \notin \Lambda_k(f), \end{cases}$$

the inequality in 29.2 holds for all $f \in L^p$. Furthermore $\mathcal{K}$ is the unique map $L^p(\mu_2) \to L^r(\mu_1)$ that equals equals $\mathcal{K}$ on $L^p \cap L^{q'}$ and is continuous.

Proof. If integrate with respect to μ_1, get something finite, so for almost every x_1 the integral is finite.

For the second part note $L^p \cap L^{q'}$ is a dense subset of L^p. □

Definition 30.2: Suppose $g \in L^q(\lambda_{\mathbb{R}^N})$ and let $K(x, y) = g(x - y)$. (Then $M_1 = M_2 = \|g\|_q$.) Let

$$\Lambda(f, g) = \left\{ \int |f(y)||g(x - y)| \, dy \right\}.$$

(Note $\Lambda(f, g)^c$ has measure 0.) Define the **convolution** of f and g to be

$$(f * g)(x) = \begin{cases} \int f(y) g(x - y) \, dy & x \in \Lambda(f, g) \\ 0, & x \notin \Lambda(f, g). \end{cases}$$

Note $*$ is commutative ($f * g = g * f$) and bilinear.

Theorem 30.3 (Young's inequality)**:** For $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$, $p, q \in [1, \infty]$,

$$\|f * g\|_r \leq \|f\|_p \|g\|_q .$$

There are two cases where this is particularly useful.

One, $p = q'$ and $r = \infty$. The inequality says

$$\|f * g\|_\infty \leq \|f\|_p \|g\|_{p'}$$

which is Hölder's. Convolution gives a continuous function. Stroock: "if you go to the aquarium they have spines and if they make love they have to rub off some of the spines. Convolution is like rubbing, scraping off the irregularities."

Take $p \in [1, \infty)$. Letting $T_h(y) = y + h$, and letting $T_f h := f \circ T_h$.

Lemma 30.4: Then

$$\lim_{h \to \infty} \|T_h f - f\|_p = 0.$$

Proof. First think of a set of f's which are dense in L^p for which this set is true is trivial. Then show the set of f such that the result holds is closed.

Thus it suffices to show this for continuous functions with compact support (dense in L^p). The desired set is closed because if $f_n \to f$ in L^p then

$$\|T_h f - f\|_p \leq \|T_h(f - f_n)\|_p + \|T_h f_n - f_n\|_p + \|f_n - f\|_p$$

by the triangle inequality and Minkowski's inequality. Taking the lim sup,

$$\begin{aligned} \overline{\lim_{|h| \to 0}} \|T_h f - f\|_p &\leq \overline{\lim_{|h| \to 0}} \|T_h(f - f_n)\|_p + \overline{\lim_{|h| \to 0}} \|T_h f_n - f_n\|_p + \overline{\lim_{|h| \to 0}} \|f_n - f\|_p \\ &\leq 2 \|f - f_n\|_p + 0 + \limsup \|f_n - f\|_p . \end{aligned}$$

For $p < \infty$ BLAH DARN YOU ERASED THIS. If $p = \infty$, $p' = \infty$, do the same thing but put on g instead of n.

Vitalle's Lemma says if $\lambda_{\mathbb{R}}(\Gamma) \in (0, y)$, then 1_Γ? $\Gamma - \Gamma \supseteq (-\delta, \delta)$. We know

$$u = 1_\Gamma \times 1_{-\Gamma}$$

If $u > 0$ then $u \in \Gamma - \Gamma$. □

When we convolve we get smoother objects!

The other common case is when $q = 1, p = r$. Then $f \in L^p$ and $g \in L^1$ give $f * g \in L^p$.

Theorem 30.5: Suppose $p = q'$ and $g \in C^1$, with $g \in L^q$ and $\frac{\partial g}{\partial x_1} \in L^q$. Take $f \in L^p$. Then $\partial_{x_i}(f * g) = f * (\partial_{x_i} g)$.

Proof. Write out difference quotients... □

Taking $g \in C^\infty$, we can keep on passing derivatives through the convolution. Writing $\partial^\alpha = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n}$ where $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^N$, we have $\partial^\alpha(f * g) = f * (\partial^\alpha g)$.

Theorem 30.6: For $g \in L^1(\lambda_N)$, define $g_t(x) = t^{-N} g(t^{-1}x)$. (We scale in a way preserving the total interval; for small t it concentrates around the origin.) Given $\int g = 1$ and $f \in L^p$, $f * g_t \to f$ in L^p as $t \searrow 0$.

Proof. Making a change of variables,

$$f * g_t(x) = \int (f(x-y) - f(x)) g_t(y)\, dy = \int (f(x-ty) - f(x)) g(y)\, dy.$$

Taking the L_p norm as a function of x,

$$\|f * g_t(x)\|_p = \left\| \int (f(x-ty) - f(x)) g(y)\, dy \right\|_p = \|(f(x-ty) - f(x)) g(y)\|_{(1,p)}.$$

But we know

$$\|T_{-ty} f - f\|_p$$

from the previous lemma, so combining these two we get a way of approximating an arbitrary element of L^p by a smooth function. □

We want $g \in C_c^\infty(B(0,1), [0,\infty))$ such that $\int g = 1$. Define

$$\Psi(t) = \begin{cases} e^{-\frac{1}{t}}, & t \ge 0 \\ 0, & t < 0; \end{cases}$$

this is a C^∞. Now consider

$$\Psi(t) = \begin{cases} e^{-\frac{1}{1-|x|^2}}, & |x| < 1 \\ 0, & t \ge 1; \end{cases}$$

and normalize it so it has integral 1.

We want $\eta \in C^\infty$ with range in $[0,1]$ such that $1_\Gamma \le \eta \le 1 \le 1_{\Gamma(\varepsilon)}$. We can let $\eta = g_\varepsilon * 1_\Gamma$.

Lecture 31
Fri. 4/22/11

§1 Hilbert spaces

Consider $L^2(\mu; \mathbb{R})$, with norm

$$\|f\|_2 = \left(\int |f|^2\, d\mu \right)^{\frac{1}{2}}.$$

2 is a particularly nice power to raise a number to, because there is a natural inner product

$$\langle f, g \rangle_{L^2(\mu,\mathbb{R})} = \int fg\, d\mu.$$

(If $E = \{1, \dots, n\}$ and $\mu(\{i\}) = 1$ then $L^2(\mu; \mathbb{R}) \cong \mathbb{R}^n$ naturally: $\langle f, g \rangle_{L^2(\mu,\mathbb{R})} = \sum_{i=1}^n f(i)g(i)$ and $\|f\|_{L^2(\mu;\mathbb{R})} = (\sum_{i=1}^n f(i))^{\frac{1}{2}}$.)

By Cauchy-Schwarz's inequality

$$|\langle f, g \rangle| \leq \|f\|_2 \|g\|_2$$

This leads to the Triangle Inequality

$$\|f + g\|_2 \leq \|f\|_2 + \|g\|_2 .$$

Definition 31.1: H is a Hilbert space over $\mathbb{R}$ if H is a vector space over $\mathbb{R}$ with an inner product making H a complete metric space.

I.e. there exists a bilinear map $(x, y) \in H^2 \to \langle x, y \rangle_H \in \mathbb{R}$ so that

1. For all y, $x \to \langle x, y \rangle_H$ is linear and symmetric ($\langle x, y \rangle_H = \langle y, x \rangle$).
2. $\langle x, x \rangle_H \geq 0$; equality holds iff $x = 0$.
3. If $\|x\|_H = \sqrt{\langle x, x \rangle}$ (which is a valid norm by Cauchy-Schwarz $\implies$ triangle inequality), then the metric determined by $\|\cdot\|_H$ is a complete metric space.

Over the complex numbers, replace $\mathbb{R}$ and $\mathbb{C}$ and symmetric by Hermitian ($\langle x, y \rangle = \overline{\langle y, x \rangle}$).

To prove Cauchy-Schwarz in the complex case, use

$$0 \leq \left\| tx + \frac{1}{t} y \right\|^2 = t^2 \|x\|_H \pm 2\Re\langle x, y \rangle + \frac{1}{t^2} \|y\|^2 .$$

Choose θ with $|\theta| = 1$ so that $\overline{\theta}\langle x, y \rangle = \langle x, \theta y \rangle > 0$.

The primary example is $L^2(\lambda_{[0,1]}, \mathbb{R})$. $C([0, 1]; \mathbb{R})$ is dense, but not closed because not every element of L^2 is continuous.

Let L be a closed linear subspace of H.

Lemma 31.2: For every $x \in H$ there exists a unique $\Pi_L x \in L$ such that $\|x - \Pi_L x\|_H = \min\{\|x - y\|_H : y \in L\}$.

Proof. By Pythagorean this is equivalent to finding $\Pi_L x \in L$ such that $x - \Pi_L x \perp L$ (i.e. $\langle x - \Pi_L x, y \rangle = 0$ for all $y \in L$): Suppose $y_0 \in L$ such that $\|x - y_0\| = \min_{y \in L} \|x - y\|_H$. Now $\|x - y_0 + ty\|^2$ has a minimum at $t = 0$ and equals $\|x - y_0\|^2 + 2t\Re\langle x - y_0, y \rangle + t^2 \|y\|^2$. Use the same trick to get rid of the real part; differentiate and use the first derivative test.

The second formulation makes it clear that the point is unique (if there are 2 points y_1, y_2, then $x - y_1, x - y_2 \perp L$, and $y_1 - y_2 \perp L$ and is in L, so equals 0).

For existence, take y_n so that $\|x - y_n\|$ tends to the minimum distance. We use the parallelogram law

$$\|a + b\|^2 + \|a - b\|^2 = 2 \|a\|^2 + 2 \|b\|^2 .$$

We show the y_n are a Cauchy sequence. Take $a = x - y_m$ and $b = x - y_n$ to get

$$\begin{aligned} \|2x - y_m - y_n\|^2 + \|y_m - y_n\|^2 &= 2\|x - y_m\|^2 + 2\|x - y_n\|^2 \\ 4\left\|x - \frac{y_m + y_n}{2}\right\|^2 + \|y_m - y_n\|^2 &= 2\|x - y_m\|^2 + 2\|x - y_n\|^2 \\ \|y_m - y_n\|^2 &\le 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - 4\delta^2 \to 0. \end{aligned}$$

Thus by completeness it converges to the point at minimum distance. This proves the generalization that given a closed convex subset of a Hilbert space, there exists a unique point at closest distance. (We only need to know $\frac{y_m+y_n}{2} \in L$ which is also true of a convex set.) □

Lecture 32
Mon. 4/25/11

§1 Hilbert spaces

Theorem 32.1: Let H be a Hilbert space and let $S \subseteq H$. Then $\text{span}(S)$ is dense in H iff $S^\perp = \{0\}$.

Proof. First suppose $\text{span}(S)$ is dense in H. If $x \perp S$ then $x \perp \text{span}(S)$ and $x \perp \overline{\text{span}(S)}$. The last step in more detail: Suppose $x \perp \Gamma$, and $y_n \to y$. This means $\langle x, y_n \rangle = 0$. By Schwarz's inequality the inner product is continuous, so $\langle x, y_n \rangle \to \langle x, y \rangle = 0$. Now $x \perp \overline{\text{span}(S)} = H$ gives $x = 0$.

For the inverse, note that $S^\perp$ is always a closed subspace, by continuity of inner product. Suppose $\text{span}(S)$ is not dense. Then there exists $x \notin \overline{\text{span}(S)}$ and $x - \Pi_{\overline{\text{span}(S)}} x \neq 0$, $x - \Pi_{\overline{\text{span}(S)}} \in \perp S \neq \{0\}$. □

Definition 32.2: $B \subseteq H$ is a basis if $\overline{\text{span}(B)} = H$ and B is linearly independent.

This agrees with the usual definition for finite dimensional H. Consider $l^2(\mathbb{N}; \mathbb{R})$, i.e. $(\alpha_1, \alpha_2, \ldots)$ such that $\sum_{i=1}^n \alpha_i^2 < \infty$. Define e_n by $(e_n)_i = \delta_{in}$. The span of the e_n are those vectors with a finite number of nonzero entries; we have to take the closure.

We care about bases with the additional property that they are orthonormal.

Lemma 32.3: Suppose that H is a separable Hilbert space. Then there exists an orthonormal basis $\{e_n | n \ge 0\}$ for H.

Proof. Choose $\{x_k : k \ge 0\}$ dense in H, and weed out the ones that are linearly dependent with the ones before: take x_{n_0} to be the first nonzero vector. Given $x_{n_0}, \ldots, x_{n_k}$, find the first $n_{k+1} > n_k$ so that x_{n_k} is linearly independent from $x_{n_0}, \ldots, x_{n_k}$. In this way we get a sequence $\{y_n | n \ge 0\}$ that are linearly dependent and that spans the same set as the x_n, so its span is dense.

Now we turn the y_n into an orthonormal basis by Gram-Schmidt orthogonalization. Consider $L_n = \text{span}(\{y_0, \ldots, y_n\})$. We claim that all finite-dimensional vector spaces are closed. Indeed, suppose $x_k = \sum_{m=1}^n \alpha_{m,k} y_m \to x$. There is $\varepsilon > 0$ such that $\|\sum_{m=1}^n \alpha_m y_m\| \geq \varepsilon \left(\sum_{m=1}^n |\alpha_m|^2\right)^{\frac{1}{2}}$: look at it as a function on the unit sphere in n dimensions; by compactness there is a minimum ε. Now $\sum_{m=1}^n (\alpha_{m,l} - \alpha_{m,k})^2 \to 0$ so the $\alpha m, l$'s converge by Cauchy.

Now take $e_0 = \frac{y_0}{\|y_0\|}$. Now take

$$e_n = \frac{y_n - \Pi_{L_{n-1}} y_n}{\left\|y_n - \Pi_{L_{n-1}} y_n\right\|}$$

to get an orthonormal basis. (The denominators are not 0 by linear independence.) Indeed, $\text{span}(\{y_0, \ldots, y_n\}) = \text{span}(\{e_0, \ldots, e_n\})$, so the span of the e_n's equals the span of the y_n's, which is dense. □

Note an orthonormal set $\{e_n\}$ is linearly independent, by taking inner products with e_m.

Theorem 32.4: Let $\{e_n : n \geq 0\}$ be an orthonormal sequence.

1. If $\{\alpha_m | m \geq 0\} \in \ell^2(N, \mathbb{C})$, then $\sum_{m=0}^n \alpha_m e_m$ converges.

2. Moreover,
$$\langle \sum_{m=0}^{\infty} \alpha_m e_m, \sum_{m=0}^{\infty} \beta_m e_m \rangle = \sum_{m=0}^{\infty} \alpha_m \overline{\beta_m}.$$

Therefore $(\alpha_m) \to \sum_{m=0}^n \alpha_m e_m$ is a isomorphism of Hilbert spaces.

Proof. 1. Use Cauchy's criterion. We have

$$\left\|\sum_{m=0}^{N} \alpha_m e_m - \sum_{m=0}^{N} \alpha_m e_m\right\|^2 = \left\|\sum_{m=M+1}^{N} \alpha_m e_m\right\|^2 = \sum_{m=M+1}^{N} |\alpha_m|^2 \to 0$$

showing this is a Cauchy sequence.

2. Similar. Do for finite, pass limits.

□

Lecture 33
Wed. 4/27/11

§1 Fourier Series

Consider the Hilbert space $L^2(\lambda_{[0,1]}, \mathbb{C})$.

Theorem 33.1 (Fourier series)**:** Define

$$e_m(x) = e^{2\pi inx}.$$

Then $\{e_n \mid n \in \mathbb{Z}\}$ is an orthonormal basis for $L^2(\lambda_{[0,1]}, \mathbb{C})$.

We give two proofs. From integration $\langle e_m, e_n\rangle = \delta_{m,n}$. We need to show that $\overline{\operatorname{span}(\{e_n : n \in \mathbb{Z}\})}$.

Proof. To do this we use Theorem 32.1: $S \subseteq H$ is dense iff $S^\perp = \{0\}$. So it suffices to show that if $f \in L^2(\lambda_{[0,1]}, \mathbb{C})$ is continuous and perpendicular to all the e_n, then $f = 0$, since the set of continuous functions is dense. We may further assume f is periodic, because we can change it on a set of small measure so that it is.

There is a 1-to-1 correspondence between periodic functions on $[0, 1]$ (i.e. with $f(0) = f(1)$) and functions on $S^1\left(\frac{1}{2\pi}\right)$ (the circle with radius $\frac{1}{2\pi}$). Thinking of this circle as on the complex plane, the e_n become the functions z^n. Now we need to show if we have a function continuous on the circle, and is orthogonal to all the z^n, then it must be 0. But the linear span of the z^n is the set of polynomials, which is dense in the set of continuous functions by the Stone-Weierstrass Theorem. □

Proof. Let

$$P(r, x) = \sum_{n\in\mathbb{Z}} r^{|n|} e_n(x), \quad r \in [0, 1), x \in [0, 1].$$

Using the geometric series formula gives

$$P(r, x) = \frac{1}{1 - re_1(x)} + \frac{r}{1 - re_{-1}(x)} = \frac{1 - r^2}{|1 - re_{-1}(x)|^2} > 0.$$

Now for $\varphi \in C^1([0, 1], \mathbb{C})$, define

$$u_\varphi(r, x) = \int_{[0,1]} P(r, x - y)\varphi(y)\, dy.$$

We claim that

$$u_\varphi(r, x) \to \varphi(x)$$

in a uniformly bounded manner. This will mean that any function orthogonal to all e_n has to be orthogonal to u_φ for every continuous periodic function φ and every $r \in [0, 1)$. Then that function is perpendicular to every φ, and we will be done.

Now

$$\int_{[0,1]} P(r, x)\, dx = 1$$

for all $r \in (0, 1)$, so $P(r, x)$. Note $\|u_\varphi\|_\infty < \|\varphi\|_\infty$. Now

$$u_\varphi(x) - \varphi(x) = \int_{[0,1]} P(r, x - y)(\varphi(y) - \varphi(x))\, dy \to 0$$

because φ is continuous. (Break up into $|x - y| < \delta$ and $|x - y| \geq \delta$ to see.) □

Corollary 33.2: The set

$$\{\cos 2\pi nx : n \geq 0\} \cup \{\sin 2\pi nx : n \geq 1\}$$

is an orthonormal basis for $L^2(\lambda_{[0,1]}, \mathbb{R})$.

Proof. If a function is perpendicular to all the functions above then considering it as a function of $\mathbb{C}$ it is perpendicular to all the e_n, so is 0. □

Corollary 33.3: For $\varphi \in L^2(\lambda_{[0,1]}, \mathbb{R})$,

$$\varphi = \sum_{n \in \mathbb{Z}} \langle \varphi, e_n \rangle e_n(x)$$

in the sense of L^2. Hence the partial sums tend to φ in measure.

Warning: We don't necessarily have pointwise convergence.

If we impose regularity conditions on the function, though, then can easily prove nice convergence. We take advantage of this by integration by parts—bring out properties by differentiation. Suppose $\varphi \in C^1([0,1], \mathbb{C})$. For $n \neq 0$, using integration by parts, (there are no boundary terms by periodicity)

$$\langle \varphi, e_n \rangle = (2\pi i n)^{-1} \langle \varphi', e_n \rangle$$

(the factor out front comes from the integral of e_n) which says that φ is converging absolutely:

$$\sum_{|n| \geq m} |\langle \varphi, e_n \rangle| \leq \frac{1}{2\pi} \sum_{|n| \geq m} \frac{1}{|n|} |\langle \varphi', e_n \rangle| \leq \frac{1}{2\pi} \left(\sum_{|n| \geq m} \frac{1}{n^2} \right)^{\frac{1}{2}} \|\varphi'\|$$

where we used Schwarz's inequality.

Theorem 33.4: Of $\{e_n\}$ is a orthonormal basis for $L^2(\mu, \mathbb{C})$ and $\{f_n\}$ is an orthonormal basis for $L^2(\nu, \mathbb{C})$ then $\{e_m f_n\}$ is an orthonormal basis for $L^2(\mu \times \nu, \mathbb{C})$.

Hence $\{e_m(x) e_n(y)\}$ is an orthonormal basis for $L^2(\lambda_{[0,1]^2}, \mathbb{C})$.

We want to know whether there is almost everywhere convergence now. The functions are now indexed by (m, n). If we take partial sum to be $|m| \leq N$ and $|n| \leq N$ we do have almost everywhere convergence (Fefferman, using a Fubini argument). However Carleson's theorem is false if you take partial sums as $(m^2 + n^2)^{\frac{1}{2}} \leq N$.

Lecture 34
Fri. 4/29/11

§1 Fourier Transform, L^1

For $x \in \mathbb{R}^N$ define

$$e_\xi(x) = e^{2\pi i \langle \xi, x \rangle}.$$

Definition 34.1: The **Fourier transform** of f is

$$\hat{f}(\xi) = \int_{\mathbb{R}^N} f(x) e_\xi(x)\, dx.$$

Define

$$\check{f}(\xi) = \hat{f}(-\xi).$$

This is well-defined for $f \in L^1(\lambda_{\mathbb{R}^N})$. Note the contrast with Fourier series: For Fourier series we can put in any measurable function, but for Fourier transform we need integrability. We will later extend the definition to functions in L^2 through a limiting procedure (through functions in $L^1 \cap L^2$).

First, note $\hat{f}$ is continuous by Lebesgue dominated convergence. (The integrand is uniformly bounded and e_ξ is continuous in ξ.) We have $\left\|\hat{f}\right\|_\infty \le \|f\|_1$.

We claim $\hat{f} \in C_0(\mathbb{R}^N, \mathbb{C})$, i.e. $\lim_{|x|\to\infty} \hat{f}(x) = 0$. It suffices to prove this for a dense subset of L^1 (continuous functions of compact support) because the set of functions vanishing at ∞ is closed. By integration by parts,

$$\begin{aligned}\widehat{\Delta f}(\xi) &= \int_{\mathbb{R}^N} \Delta f(x) e_\xi(x)\, dx \\ &= -(2\pi|\xi|)^2 \hat{f}(\xi)\end{aligned}$$

which tends to 0 at ∞, using $\hat{f}$ is bounded (because it is continuous with compact support).

To what extent can we recover f from $\hat{f}$? Does $\hat{f}$ uniquely determine f; i.e. if $\hat{f}$ vanishes everywhere does f vanish a.e.? Defining $P(r,x) = \sum_{n\in\mathbb{Z}} r^{|n|} e_n(x)$, we have

$$\langle P(r,\mathbf{x}), e_n\rangle = r^{|n|}.$$

Rescale $g(x) = e^{-\pi|x|^2}$, defining

$$g_t(x) = t^{-\frac{N}{2}} e^{-\frac{\pi|x|^2}{t}}.$$

Noting $\int \frac{1}{(2\pi)^{\frac{1}{2}}} e^{-\frac{x^2}{2}}\, dx = 1$, by calculation (Fubini)

$$\int_{\mathbb{R}^N} g_t(x)\, dx = 1.$$

Now let $\zeta = (\zeta_1, \ldots, \zeta_N)$ and note

$$t^{-\frac{1}{2}} \int_{\mathbb{R}} e^{2\pi\xi x} e^{-\pi x^2 t}\, dx = e^{t\pi\xi^2}.$$

(Proof: It suffices to show this for real ξ by the fact that complex analytic functions agreeing on $\mathbb{R}$ agree everywhere. Just complete the square.) Then $\hat{g}_t(\xi) = e^{-t\pi|\xi|^2} = t^{-\frac{1}{2}} g_{\frac{1}{t}}(\xi)$. Now

$$(\hat{g}_t)^\vee(x) = t^{-\frac{1}{2}} \hat{g}_{\frac{1}{t}}(-x) = t^{-\frac{1}{2}} e^{-t\pi|x|^2} = g_t(x).$$

Now consider

$$\int \hat{f}(\xi)\overline{g(\xi)}\,d\xi = \iint f(x)\overline{e_{-x}(\xi)g(\xi)}\,dx = \int f(x)\overline{\check{g}(x)}\,dx.$$

Since $\hat{f}(\xi)e_{-x}(\xi) = \widehat{f \circ T_x}(\xi)$,

$$\begin{aligned}
\int \hat{f}(\xi)e_{-x}(\xi)e^{-t\pi|\xi|^2}\,d\xi &= \int \widehat{f \circ T_x}\overline{\check{g}_t(\xi)}\,d\xi \\
&= \int f \circ T_x(y) g_t(y)\,dy \\
&= \int f(x+y)g_t(y)\,dy \\
&= \int f(x-y)g_t(y)\,dy \\
&= f * g_t(x) \xrightarrow{L^1} f.
\end{aligned}$$

Hence the Fourier transform is one-to-one.

Now $\{\hat{f} : f \in L^1\}$ contains only continuous functions vanishing at ∞, but not every continuous function vanishing at ∞ is the Fourier transform of a function. There is no good characterization of which ones are. We do know this set is an algebra: We can take products, because $\widehat{f * g} = \hat{f}\hat{g}$.

§2 Fourier Transform, L^2

We want to continuously extend the Fourier transform to L^2, i.e. $\mathcal{F} : L^2 \to L^2$ such that $\mathcal{F}f = \hat{f}$ if $f \in L^1 \cap L^2$. Note this extension will be unique if it exists because $L^1 \cap L^2$ is dense in L^1. We need to show that $\left\|\hat{f}\right\|_2 \le C\|f\|_2$; indeed we will show that $C = 1$ works. Then given $f \in L^2$, consider $1_{B(0,R)}f$. We have $\left\|1_{B(0,R_2)}f - 1_{B(0,R_1)}f\right\|_2 = \left(\int_{B(0,R_2)\setminus B(0,R_1)}|f|^2\right)^{\frac{1}{2}}$ and use the above estimate.

To show $\left\|\hat{f}\right\|_2 \le C\|f\|_2$, we will diagonalize $\mathcal{F}$, i.e. find an orthonormal basis ($\langle \tilde{h}_m, \tilde{h}_n\rangle = \delta_{m,n}$) so that $\tilde{h}_n \in L^1 \cap L^2$ and $\hat{h}_n = i^n\tilde{h}_n$—the Hermite functions.

Lecture 35
Mon. 5/2/11

§1 Fourier transform on L^2

Definition 35.1: Define the **Hermite polynomials** by

$$H_n(x) = (-1)^n e^{2\pi x^2}\frac{\partial^n}{\partial x^n}e^{-2\pi x^2}.$$

Define the **Hermite functions** by

$$h_n(x) = e^{-\pi x^2} H_n(x).$$

Note that H_n is a nth degree polynomial with leading coefficient $(4\pi)^n$. Hence $\operatorname{span}(\{H_m \mid 0 \leq m \leq n\}) = \operatorname{span}(\{x^m \mid 0 \leq m \leq n\})$. Note $h_n(x)$ is integrable.

Lemma 35.2: The h_n form an orthogonal basis for $L^2(\mathbb{R})$. Indeed,

$$\langle h_m, h_n \rangle = 2^{-\frac{1}{2}}(4\pi)^n n! \delta_{m,n}.$$

Proof. Define the **raising operator** and the **lowering operator** by

$$\begin{aligned}
a_+\varphi(x) &= 2\pi x\varphi(x) - \frac{\partial}{\partial x}\varphi(x) = -e^{-\pi x^2}\frac{\partial}{\partial x}(e^{-\pi x^2}\varphi(x)) \\
a_-\varphi(x) &= 2\pi x\varphi(x) + \frac{\partial}{\partial x}\varphi(x).
\end{aligned}$$

Note in particular

$$a_+ h_n = h_{n+1} \tag{15}$$

$$a_- h_n(x) = e^{-\pi x^2}\frac{\partial}{\partial x} H_n \in \operatorname{span}(\{h_0, \ldots, h_{n-1}\}). \tag{16}$$

Given $\varphi, \psi \in C^1$ and $(a_+\varphi)\psi, \varphi(a_-\psi) \in L^1(\lambda_{\mathbb{R}}; \mathbb{C})$. Then

$$\int (a_+\varphi)\psi = \int \varphi(a_-\psi).$$

If one of φ, ψ has compact support, then integration by parts shows the above. In general, choose a bump function $\eta \in C_c^\infty(\mathbb{R}; [0,1])$ such that $\eta(x) = 1$ for $x \in [-1,1]$ and $\eta(c) = 0$ for $x \notin [-2,2]$. Let $\eta_R(x) = \eta(R^{-1}x)$. Replace φ with η_R and take the limit as $R \to \infty$. (Note $\frac{\partial}{\partial x}(\eta_R\varphi) = \frac{1}{R}\eta'(R^{-1}x)\varphi + \eta_R\frac{\partial}{\partial x}\varphi$.)

Suppose $m \leq n$. By (15), $h_n = a_+^n h_0$, so

$$\langle h_m, h_n \rangle = \langle h_m, a_+^n h_0 \rangle = \langle a_-^n h_m, h_n \rangle.$$

If $n > m$ then $a_-^n h = 0$. If $m = n$ then we use $a_-^n h_n = (4\pi)^n n!$ and $\langle h_0, h_0 \rangle = 2^{-\frac{1}{2}}$ to get

$$\|h_n\|^2 = (4\pi)^n n! 2^{-\frac{1}{2}}.$$

□

Lemma 35.3:

$$\alpha_- h_n(x) = 4\pi n h_{n-1}.$$

Proof. We know that

$$a_- h_n = \sum_{m=0}^{n-1} \alpha_m h_m.$$

Taking inner products of h_k and using orthogonality we get

$$\langle h_k, a_- h_n \rangle = \alpha_k (4\pi)^k k! 2^{-\frac{1}{2}}.$$

But

$$\langle h_k, a_- h_n \rangle = \langle h_{k+1}, h_n \rangle = (4\pi)^n n! 2^{-\frac{1}{2}} \delta_{k,n-1}.$$

□

Now we have

$$a_+ a_- h_n = a_+(4\pi n h_{n-1}) = 4\pi n h_n.$$

By definition the LHS is also

$$a_+ a_- h_n = (4\pi x)^2 h_n - \frac{\partial^2}{\partial x^2} h_n - 2\pi h_n$$

Hence

$$(4\pi x)^2 h_n - \frac{\partial^2}{\partial x^2} = 2\pi(2n+1) h_n.$$

Adding the following two equations,

$$\begin{aligned} 2\pi x h_n - \frac{\partial}{\partial x} h_n &= h_{n+1} \\ 2\pi x h_n + \frac{\partial}{\partial x} h_n &= 4\pi n h_{n-1} \\ 4\pi x h_n &= h_{n+1} + 4\pi n h_{n-1} \\ (4\pi)^2 \int x^2 h_n^2 &= (4\pi)^{n+1} (n+1)! 2^{-\frac{1}{2}} + (4\pi n)^2 (4\pi)^{n-1} (n-1)! 2^{-\frac{1}{2}}. \end{aligned}$$

Let

$$\tilde{h}_n = \frac{2^{\frac{1}{4}} h_n}{((4\pi)^n n!)}^{\frac{1}{2}}.$$

Then the $\tilde{h}_n$ is an orthonormal basis, and

$$\int x^2 \tilde{h}_n^2 = \frac{2n+1}{4\pi}.$$

Now

$$\int (1+x^2) \tilde{h}_n^2 = 1 + \frac{2n+1}{4\pi}$$

so using Schwarz's inequality

$$\left(\int \left|\tilde{f}_n\right|\right)^2 = \left(\int (1+x^2)^{\frac{1}{2}} \frac{\tilde{h}_n}{(1+x^2)^{\frac{1}{2}}}\right) \leq ?$$

Now back to Frouier transform.
We show

$$\hat{h}(\xi) = i^n h_m(\xi).$$

Indeed,

$$(a_-)_x e_\xi(x) = i(a_+)_\xi e_\xi(x).$$

Then

$$(2\pi x + i2\pi i\xi)e_\xi(x) = i(2\pi\xi - 2\pi i x)e_\xi(x).$$

But

$$\begin{aligned}\int e_\xi(x) h_{n+1}(x)\,dx = \int e_\xi(x)(a_+)_x h_n(x)\,dx \\ = \int ((a_-)_x e_\xi(x)) h_m(x)\,dx \\ h_{n+1}(\xi) = i\int (a_+)_\xi e_\xi(x).\end{aligned}$$

Define the Fourier operator

$$\mathcal{F} = \sum_{n=0}^{\infty} i^n \langle f, \tilde{h}_n\rangle \tilde{h}_n.$$

Then $\|\mathcal{F}d\| = \|f\|$. For $f \in L^1 \cap L^2$, then this agrees with the original definition.

Lecture 36
Wed. 5/4/11

§1

Recall $\{h_n : n \geq 0\}$ is a basis. Given $f \in L^2$ with $f \perp h_n$ for all n, i.e.

$$\int \varphi(x) x^n\,dx = 0$$

for all n. Write $\varphi = fh_0$. Summing gives

$$\sum_{n=0}^{N} \frac{(2\pi i)^n}{n!} \int x^n \varphi(x)\,dx.$$

We will let $N \to \infty$ to get

$$\hat{\varphi} = \int e_\xi x \varphi(x)\, dx = 0.$$

However, we need to know that we can pass the limit, i.e.

$$\varphi(x) \sum_{n=0}^{N} \frac{(2\pi i \xi)^n}{n!} x^n \to \xi(x) e_\xi(x) \text{ in } L^1.$$

We produce a Lebesgue dominating function:

$$\langle \varphi(x) \sum_{n=0}^{N} \frac{(2\pi i \xi)^n}{n!} x^n \rangle \leq |\xi(x)|(e^{2\pi\xi x} + e^{-2\pi \xi x})^2.$$

The latter is integrable, by Schwarz's inequality:

$$\int |\xi(x)|(e^{2\pi\xi x} + e^{-2\pi\xi x})^2 \leq \|f\|_2^2 \int h_0(x)^2 (e^{2\pi\xi x} + e^{-2\pi\xi x}).$$

For any function f we know we can write it as

$$f = \sum_{n=0}^{\infty} \langle f, \tilde{h}_n \rangle_2 \tilde{h}_n.$$

Since $\hat{h_n} = i^n h_n$, we will define

$$\mathcal{F} f = \sum_{n=0}^{\infty} i^n \langle f, \tilde{h}_n \rangle_2 \tilde{h}_n.$$

Indeed $\|\mathcal{F} f\|_2 = \|f\|_2$. The remaining question is whether this definition of $\mathcal{F} f = \hat{f}$ for $f \in L^1 \mathcal{L}^2$.

It suffices to prove this for all continuous functions with compact support. Indeed, if $f \in L^p \cap L^q$ we can approximate f with continuous functions with compact support tending to both f in both L^p and L^q.

Consider $f_n = \sum_{m=0}^{n} \langle f, \tilde{h}_n \rangle_2 \tilde{h}_n$. We have $\hat{f}_n = \sum_{m=0}^{n} i^n \langle f, \tilde{h}_n \rangle_2 \tilde{h}_n$. We need to show $f_n = \sum_{m=0}^{n} \langle f, \tilde{h}_n \rangle \tilde{h}_n$ with $f_n \to f$ in L^1. Now $\left\|\tilde{h}_n\right\|_1 \leq C(1+n)^{-\frac{1}{2}}$. Now

$$(2\pi x)^2 \tilde{h_n} - \frac{\partial^2}{\partial x^2} \tilde{h}_n = 2\pi(2n+1)\tilde{}$$

Let

$$\begin{aligned}
\mathcal{H}^k h_n &= (2\pi(2n+1))^{\frac{k}{2}} \tilde{h_n} \\
\tilde{h}_n &= -\frac{\mathcal{H}^k h_n t}{(2\pi(2n+1))^{\frac{k}{2}})} \\
&=
\end{aligned}$$

BLAH BLAH

§2 Radon-Nikodym

Recall that given the distribution F corresponding to the measure μ, decomposing F into an absolutely part and singular part $F = F_a + F_s$ gives a decomposition of F as $\mu_F = \mu_a + \mu_s$. Moreover, $\mu_a(\Gamma) = \int_\Gamma f\, d\lambda_{\mathbb{R}}$.

We generalize this as follows. Given $(E, \mathcal{B})$ with μ finite and ν σ-finite, we show μ can be decomposed uniquely as $\mu = \mu_a + \mu_s$ (Lebesgue decomposition) where $\mu_a(\Gamma) = \int_\Gamma f\, d\nu$.

f is called the **Radon-Nikodym form**. We want to "project μ onto ν and wee what we get

Theorem 36.1: Suppose $L : H \to G$ is a linear function. Then L is continuous iff there exists $g \in H$ such that $\Lambda(h) = \langle h \rangle = \langle h, g \rangle$

Proof. $L(x) = \langle x, \theta \rangle$. Consider, $\ker(\Gamma \mid \Lambda(h) = 0)$ □

For $L = H$, take h go be i, and use the Suppose $L \not\subseteq H$Choose $f \notin L$ and consider $\varphi \frac{f - \Pi_L f}{\|f - \Pi_n\|}$.

Blah

$$h - \frac{G(h)\varphi}{G(h)}.$$

$$\langle h, \varphi \rangle_H = \frac{\Gamma(h)}{\Gamma(\varphi)}$$

$g = \Gamma(\varphi)\varphi$.

Given any α,

$$\mu_G(\Gamma) \le L\lambda_{\mathbb{R}}(\Gamma).$$

Lemma 36.2: Suppose $\mu \le \nu$. Then there exists a unique $f \in L^1$ such that $\mu(\Gamma) = \int_\Gamma f\, d\nu$ for all $\Gamma \in \mathcal{B}$. In fact we can choose $f \in [0, 1]$.

If suffices to prove for finite (?). Consider $\Gamma(\varphi) = \int \varphi\, d\mu$. Then

$$|\Gamma(\varphi)| \le \left(\varphi^2\, d\mu\right)^{\frac{1}{2}} \mu(E)^{\frac{1}{2}}. \le \mu(E)^{\frac{1}{2}} \|\varphi\|_{L^2(\nu)}.$$

Since $\mu \le \nu$.

By Riesz Representation,

$$\Gamma(\xi) = \langle \varphi, f \rangle_{L^2(\nu)} = \int \varphi f\, d\nu = \int \varphi\, d\mu.$$

φ satisfies this.

Use

Lecture 37
Fri. 5/6/11

§1

Suppose μ is finite and ν is σ-finite with $\mu \le \nu$. Then $\int \varphi \, d\mu \le \int \varphi \, d\nu$ for $\varphi \ge 0$. We've showed that there exists $f \in L^1(\nu)$ with range in $[0,1]$ such that $\int \varphi \, d\mu = \int f\varphi \, d\nu$ for all $\varphi \in L^1$ (Riesz representation). The Lebesgue decomposition theorem says that there exists unique μ_a, μ_s such that $\mu_a \ll v$, $\mu_s \perp \nu$, and $\mu = \mu_a + \mu_s$.

Theorem 37.1 (Radon-Nikodyn)**:** There exists $f \in L^1(\nu)$ with $f \ge 0$ and $\int \varphi \, d\mu_a = \int \varphi f \, d\nu$ for all $\varphi \in L^1(\mu)$.

Definition 37.2: We call f the **Radon-Nikodyn derivative** and write

$$f = \frac{d\mu_a}{d\nu}.$$

Proof. No doubt $\mu \le \mu + \nu$. Apply the theorem (??) to find ψ such that

$$\int \varphi \, d\mu = \int \varphi\psi \, d\mu + \int \varphi\psi \, d\nu$$

$$\int \varphi(1-\psi) \, d\mu = \int \varphi\psi \, d\nu.$$

Let $B = \{\psi = 1\}$. Take $\varphi = 1_B$ to get $\nu(B) = 0$. Define $\mu_a(\Gamma) = \mu(\Gamma \cap B^c)$ and $\mu_s(\Gamma) = \mu(\Gamma \cap B)$. Clearly, $\mu_s \perp \nu$. We get

$$\int \varphi \, d\mu_a = \int_{B^c} \varphi \, d\mu = \int_{B^c} (1-\psi) \frac{\varphi}{1-\psi} \, d\mu = \int_{B^c} \varphi \frac{\psi}{1-\psi} \, d\nu$$

and $f = 1_{B^c} \frac{\psi}{1-\psi}$.

Now we prove uniqueness of f. Suppose $\mu = \mu_s + \mu_a = \mu' + \mu''$ are two representations; then $\mu_a - \mu' = \mu'' - \mu_s$. We show an absolutely continuous measure can't equal a singular measure, unless it's zero. Take A, A' such that $\nu(A) = \nu(A') = 0$ and $\mu_s(A^c) = 0$ and $\mu''(A'^c) = 0$. Take $B = A \cup A'$. Then $\nu(B) = 0$, $\mu_s(B^c) = 0$, and $\mu''(B^c) = 0$. □

At this level of generality (??) is just an existential statement. But often we want to say f is more than just L^1. Recall that $\mu \ll \nu$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\nu(\Gamma) < \delta$ implies $\mu(\Gamma) < \varepsilon$. Suppose we know that $f = \frac{d\mu}{d\nu}$ is in L^p. Then

$$\mu(\Gamma) = \int_\Gamma f \, d\nu \le \|f\|_p \, \nu(\Gamma)^{1-\frac{1}{p}}.$$

This is a much more quantitative statement.

§2 Daniell integration

Although it's intuitively appealing to start with a measure and build an integration theory, for technical reasons this isn't the most elegant way to proceed. Instead, we could start with an integration theory and figure out what measure it came from. Sets don't behave well when you put them together: they don't have a vector space structure. But functions do.

Let E be a nonempty set. Let L be a vector lattice of functions $f : E \to \mathbb{R}$. (A vector lattice is a vector space closed under the operations max and min.) Consider an "integral", a map $I : L \to \mathbb{R}$ such that

1. I is linear.
2. $I(f) \geq 0$ if $f \geq 0$.
3. If $f_n \in L$ and $f_n \searrow 0$ then $I(f_n) \searrow 0$.

For example, consider a measure space $(E, \mathcal{B}, \mu)$ with $L = L^1(\mu, \mathbb{R})$, $I(f) = \int f \, d\mu$. Property 3 is just the Monotone Convergence Theorem. (Finiteness condition holds.)

But this is not the only example... Now let $\mathcal{A}$ be an algebra of subsets, with $\mu : \mathcal{A} \to [0, \infty)$ finitely additive, so $\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2)$, $A_1 \cap A_2 = \phi$. Take $L(\mathcal{A})$ to consist of $\mathcal{A}$-measurable functions, i.e. functions with finite range, so that the inverse image of any element is in A. Define

$$I(\varphi) = \int f \, d\mu = \sum_a a\mu(\varphi = a).$$

(Proof of linearity for simple functions only used finite additivity.) However, this does not necessarily satisfy property 3. Property 3 is equivalent to: If $A_n \in \mathcal{A}$ and $A_n \searrow \phi$ then $\mu(A_n) \searrow 0$. The forward direction is from applying the property to indicator functions. For the reverse direction,

$$I(f_n) \leq \varepsilon \underbrace{\mu(f_n \leq \varepsilon)}_{\leq \varepsilon\mu(E)} + \underbrace{\|f_n\|_u \, \mu(f_n \geq \varepsilon)}_{\to 0}.$$

Consider E a compact Hausdorff space. Consider $L = C(E, \mathbb{R})$. Take I to be a nonnegative linear function on L. Then $(E, C(E; \mathbb{R}), I)$ is an integration. We check property 3: Note $|I(f)| \leq I(1) \|f\|_u$. Indeed, $\|f\|_u - f$ is nonnegative, so $I(1) \|f\|_u - I(f) \geq 0$. Similarly for the other direction. Second, note $f_n \searrow 0$ implies $\|f_n\|_u \searrow 0$, Dini's Lemma.

Lemma 37.3 (Dini)**:** Blah above.

Proof. For every $x \in E$ there exists $n(x)$ such that $f_n(x) < \varepsilon$ for all $n \geq n(x)$. Therefore, by continuity, for all $x \in E$ we can choose $U(x)$ such that $|f_{n(x)}(x) - f_{n(x)}(y)| \leq \varepsilon$ if $y \in U(x)$. Now apply Heine-Borel to see a finite number of $U(x)$ cover the space. $\square$

Lecture 38
Mon. 5/9/11

If we add the condition that the constant function 1 to our integration theory $\mathcal{I} = (E, L, I)$, then it must have come from the integral with respect to some measure. To prove this we have to construct the measure; in order to construct the measure we have to extend the definition of I.

Let E be a compact space and $L = C(E, R)$. Let I be a nonlinear linear functional. If f came from μ then $I(f) = \int f\, d\mu$. We would like to say the corresponding measure is

$$\mu(\Gamma) = I(1_\Gamma).$$

But $I(1_\Gamma)$ isn't necessarily defined, so we need to extend the definition of I.

Let L_u be the set of $f : R \to (-\infty, \infty]$ such that there exist $\varphi_n \nearrow f$ where $\{\varphi_n \mid n \geq 1\} \subseteq L$.

Claim 38.1: There is a unique extension of I to L_u such that $I(f) = \lim_{n\to\infty} I(f_n)$ if $\{f_n \mid n \geq 1\} \subseteq L_u$ and $f_n \nearrow f$.

Proof. This is analogous to the proof that we can extend a Lebesgue integral for nonnegative simple functions to nonnegative measurable functions (Lemma 13.2). Think of L as playing the role of simple functions.

The essential step was to prove that the integral of simple functions satisfied the condition that if $\varphi_n \nearrow \psi \in L$, then $\psi \leq \lim_{n\to\infty} \varphi_n$ implies $I(\psi) \leq \lim_{n\to\infty} I(\varphi_n)$. $\square$

Define

$$\overline{I}(f) = \inf\{I(\varphi) \mid \varphi \in L_u,\ \varphi \geq f\}$$
$$\underline{I}(f) = \sup\{-I(\varphi) \mid \varphi \in L_u,\ -\varphi \leq f\}$$

Note that,

$$\begin{aligned}
\overline{I}(\alpha f) &= \alpha \overline{I}(f), && \alpha \geq 0\\
\underline{I}(\alpha f) &= \alpha \underline{I}(f), && \alpha \geq 0\\
\overline{I}(\alpha f) &= \alpha \underline{I}(f), && \alpha \geq 0\\
\underline{I}(\alpha f) &= \alpha \overline{I}(f), && \alpha \leq 0\\
\overline{I}(f+g) &\leq \overline{I}(f) + \overline{I}(g)\\
\underline{I}(f+g) &\geq \underline{I}(f) + \underline{I}(g).
\end{aligned}$$

We're creeping up on linearity. If we look at f such that $\overline{I}(f) = \underline{I}(f)$ then we do in fact have linearity. Define

$$\mathfrak{M}(\mathcal{I}) = \{f \mid \overline{I}(f) = \underline{I}(f)\}$$
$$L^1(\mathcal{I}, \mathbb{R}) = \{f \in \mathfrak{M}(\mathcal{I}) : \overline{I} < \infty\}$$

and write $\tilde{I}(f) = \overline{I}(f) = \underline{I}(f)$ if $f \in L^1(\mathcal{I})$.

Theorem 38.2: $(E, L^1(\mathcal{I}, \mathbb{R}), \tilde{I})$ is again an integration theory. Moreover, $\tilde{I}|_L = I|_L$. Finally, if $\{f_n \mid n \geq 1\} \subseteq L^1(\mathcal{I}, \mathbb{R})$ and $f_n \nearrow f$, then $f \in L^1(\mathcal{I}, \mathbb{R})$ and $\tilde{I}(f_n) \nearrow \tilde{I}(f)$ if $\sup_n f_n(x) < \infty$ for all x if $\tilde{I}(f) < \infty$.

Note if our original integration theory came from a complete measure, and $L = L^1(\mu, \mathbb{R})$, then you'll end with the same thing with the construction above, by the Monotone Convergence Theorem.

When can the above be carried out? When $1 \in L$. Assume $1 \in L$. Define $\sigma(L)$ to be the smallest σ-algebra over E with respect to which every $f \in L$ is measurable:

$$\sigma(L) = \sigma(\{\{f > a\} \mid f \in L,\, a \in \mathbb{R}\}).$$

Theorem 38.3 (Stone)**:**

$$\sigma(L^1(\mathcal{I}, \mathbb{R})) = \{\Gamma \mid 1_\Gamma \in L^1(\mathcal{I}, \mathbb{R})\}.$$

Moreover, if $\mu_I(\Gamma) = \tilde{I}(1_\Gamma)$ for $\Gamma \in \sigma(L^1(\tilde{I}))$ then $\sigma(I)$ is a finite measure, $L^1(\mu_I) = L^1(\tilde{I})$, and $\tilde{I}(f) = \int f\, d\mu_I$ for $f \in L^1(\mu_I)$.

Proof. Suppose $f \in L(\tilde{I}, \mathbb{R})$. Then $1_{f>a} \in L^1(\tilde{I})$. Consider the functions (in L^1),

$$[n(f - f \wedge a)] \wedge 1 \nearrow 1_{f>a}.$$

So the indicator function is in L^1. □

Let $\mathcal{A}$ be an algebra $\mu : \mathcal{A} \to [0, \infty)$ that is finitely additive and such that $A_n \searrow \phi$ implies $\mu(A_n) \searrow 0$. Let $I(\varphi) = \int \varphi\, d\mu$. By Stone's theorem, the extension is the measure μ_I.

There is a basic deficiency. We have to have finite measure to carry this process through. (We needed that 1 is measurable.) With the usual trick, though, we can extend the result to σ-finite measures. However, it cannot be extended to wildly infinite measures.

Why would we care about wildly infinite measures? Recall the Hausdorff measure

$$H^{N,\delta}(\Gamma) = \inf \left\{ \Omega_N \sum_{C \in \mathcal{C}} \operatorname{rad}(C)^N : \|\mathcal{C}\| \leq \delta \right\} \tag{17}$$

with

$$H^N(\Gamma) = \lim_{\delta \searrow 0} H^{N,\delta}(\Gamma). \tag{18}$$

We had

$$H^N(\Gamma) = \lambda_{\mathbb{R}^N}(\Gamma) = H^{N,\delta}(\Gamma) \quad \Gamma \in \mathcal{B}_{\mathbb{R}^N}$$

The problem of Hausdorff measure is that a lot of interesting sets are invisible, for example, any hyperplane has measure 0. But Hausdorff's description of Lebesgue measure gives a generalization that make small sets look visible. Take a lower power of the radius: replace N by some $s < N$ in the definitions above. This is no longer σ-finite!

Why do we care? Because of the theory of fractals. They are calibrated using Hausdorff measure. Another motivation is that Hausdorff measure arises in geometric measure theory in differential geometry: indeed, $\lambda_M = H^{N-1}|_M$. In general, if M is an n-dimensional submanifold in $\mathbb{R}^N$, then $\lambda_M = H^n|_M$. This is a very powerful observation that allows integration theory in much greater generality than ordinary differential geometry (which has a hard time dealing with things that aren't continuously differentiable).

Lecture 39
Wed. 5/11/11

§1 Caratheodory

Definition 39.1: Suppose $E \neq \phi$. A function $\tilde{\mu} : \mathcal{P}(E) \to [0, \infty]$ is a **outer measure** if

1. $\tilde{\mu}(\phi) = 0$
2. $\tilde{\mu}(A) \le \tilde{\mu}(B)$ whenever $A \subseteq B$.
3. $\tilde{\mu}\left(\bigcup_{n=1}^{\infty} \Gamma_n\right) \le \sum_{n=1}^{\infty} \tilde{\mu}(\Gamma_n)$.

Note if $(E, \mathcal{B}, \mu)$ is an outer measure, then

$$\tilde{\mu}(\Gamma) = \inf\{\mu(B) \mid B \in \mathcal{B},\, \Gamma \subseteq B\}.$$

Definition 39.2: A set $A \subseteq \mathcal{P}(E)$ is **Caratheodory measurable** if for every $\Gamma \subseteq E$, $\tilde{\mu}(\Gamma) = \tilde{\mu}(A \cap \Gamma) + \tilde{\mu}(A^c \cap \Gamma)$, i.e. A cuts every other set nicely from the point of view of μ.

The collection of Caratheodory measurable sets is denoted $\mathcal{B}_{\tilde{\mu}}$.

Note if $\tilde{\mu}$ comes from a measure, then everything in the original σ-algebra is Caratheodory measurable.

The main theorem is the following.

Theorem 39.3: $\mathcal{B}_{\tilde{\mu}}$ is a σ-algebra and $\tilde{\mu}|_{\mathcal{B}_{\tilde{\mu}}}$ is a measure.

This theorem is very general. However, for it to be of use, we need a criteria on which sets are Caratheodory. This is given by the following.

Theorem 39.4: Let (E, ρ) be a metric space. If $\rho(A, B) > 0$ then $\tilde{\mu}(A \cup B) = \tilde{\mu}(A) + \tilde{\mu}(B)$. Then $\mathcal{B}_E \subseteq \mathcal{B}_{\tilde{\mu}}$.

Lemma 39.5: The set of Caratheodory measurable sets is an algebra.

Proof. It is clear that φ, E are measurable sets, and that A is measurable iff A^c is measurable, by symmetry of definition.

Suppose $A, B \in \mathcal{B}_{\tilde{\mu}}$. Let $\Gamma \subseteq E$. Then

$$\begin{aligned}\tilde{\mu}(\Gamma) &= \tilde{\mu}(A \cap \Gamma) + \tilde{\mu}(A^c \cap \Gamma) \\ &= \tilde{\mu}(A \cap \Gamma) + \tilde{\mu}(A^c \cap B \cap \Gamma) + \tilde{\mu}(A^c \cap B^c \cap \Gamma) \\ &= \tilde{\mu}(A \cap \Gamma) + \tilde{\mu}(A^c \cap B \cap \Gamma) + \tilde{\mu}((A \cup B)^c \cap \Gamma) \\ &\geq \tilde{\mu}((A \cup B) \cap \Gamma) + \tilde{\mu}(A \cup B)^c \cap \Gamma\end{aligned}$$

where we used... □

Lemma 39.6: The set of Caratheodory measurable sets is a σ-algebra.

Proof. It suffices to show that if $\{A_n \mid n \geq 1\} \subseteq \mathcal{B}_{\tilde{\mu}}$ are mutually disjoint, then their intersection is also in $\mathcal{B}_{\tilde{\mu}}$. Let $B_n = \bigcup_{m=1}^{n}$. By induction on n, $\tilde{\mu}(\Gamma \cap B_n) = \sum_{m=1}^{n} \tilde{\mu}(\Gamma \cap A_m)$. By monotonicity,

$$\tilde{\mu}(\Gamma \cap B) \geq \tilde{\mu}(\Gamma \cap B_n) = \sum_{m=1}^{n} \tilde{\mu}(\Gamma \cap A_m).$$

Taking the limit as $n \to \infty$ and letting $B = \bigcup_{m=1}^{\infty} A_m$,

$$\tilde{\mu}(\Gamma) = \tilde{\mu}(\Gamma \cap B_n) + \tilde{\mu}(\Gamma \cap B_n^c) \geq \sum_{m=1}^{n} \tilde{\mu}(\Gamma \cap A_m) + \tilde{\mu}(\Gamma \cap B^c) \geq \tilde{\mu}(\Gamma \cap B).$$

What?

$$\tilde{\mu}(\Gamma) \geq \sum_{m=1}^{\infty} \tilde{\mu}(\Gamma \cap A_m) + \tilde{\mu}(\Gamma \cap B^c) \geq \tilde{\mu}(\Gamma \cap B) + \tilde{\mu}(\Gamma \cap B^c)$$

$$\tilde{\mu}\left(\bigcup_{m=1}^{n} A_m\right) \geq \sum_{m=1}^{\infty} \tilde{\mu}(A_m).$$

So measure. □

Lemma 39.7: If G is open, $A \subseteq G$, and $A_n = \{x \in A \mid \rho(x, G^c) \geq \frac{1}{n}\}$, Then $\tilde{\mu}(A_n) \nearrow \tilde{\mu}(A)$.

Proof. Since G is open, the A_n increase to A. Let $D_n = A_{n+1} \backslash A_n$. Let

$$\begin{aligned}\tilde{\mu}(A) = \tilde{\mu}(A_{2m}) \cup \bigcup_{n \geq 2m} D_n &\leq \tilde{\mu}(A_{2m}) + \tilde{\mu}\left(\bigcup_{n \geq m} D_{2n}\right) + \tilde{\mu}\left(\bigcup_{n \geq m} D_{2n+1}\right) \\ &\leq \sum_{n \geq m} \tilde{\mu}(D_{2n}) + \sum_{n \geq m} \tilde{\mu}(D_{2n+1})\end{aligned}$$

We need to show that for $\mu(A_N)$,

$$\tilde{\mu}(A) \leq \lim_{n \to \infty} \tilde{\mu}_{n \to iy} \tilde{\mu}(A_n) =: K < \infty.$$

Note

$$K \geq \tilde{\mu}(A_{2M+1}) \geq \tilde{\mu}\left(\bigcup_{n=1}^{M} D_{2n}\right)$$

created space between them (think of ring of circles) Positive distance from one another. Use thm. □

Lemma 39.8:

$$\mathcal{B}_E \subseteq \mathcal{B}_{\tilde{\mu}}.$$

Proof. Let $A = G \cap \Gamma$, $A_n \nearrow A$, $A_n \nearrow A$... Now

$$\tilde{\mu}(G \cap \Gamma) = \lim_{n \to \infty} \tilde{\mu}(A_n)$$
$$\tilde{\mu}(A_a \cup G^c)\tilde{\mu}(A_n) + \tilde{\mu}(G_c).\tilde{\mu}(\Gamma) \qquad \geq \tilde{\mu}() + \tilde{\mu}()$$

$\Gamma = (\Gamma \cap G) \cup (\Gamma \cup G^c) \supseteq A_n \cup (\Gamma \cap G^c)$. A_n a positive distance from G^c. These sets a pos distance

$$\begin{aligned}\tilde{\mu}(\Gamma) &\geq \tilde{\mu}(A_n) + \tilde{\mu}(\Gamma \cap G^c)\\ &\geq \tilde{\mu}(\Gamma \cap G) + \tilde{\mu}(\Gamma \cap G^c)\end{aligned}$$

□