Intro to Modern Analysis II

Fabio Nironi

Transcribed by Ron Wu

This is an advanced undergraduate course, offered in spring 2013 at Columbia University. Course textbook is Rudin, *Principles of Mathematical Analysis.* Recommended books for measure theory are Rudin, *Real and Complex Analysis.* Royden, *Real Analysis.* Office hours: T,R 2:30-3:30.

Contents

Course Overview

Lecture 1 (1/22/13)

This course is continuation of Analysis 1, and we will cover Rudin ch 7-9, then we will do some more advanced topics: measure theory + Lebesgue Integral.

Plan

(1) Sequences of functions (chapter 7)

Uniform Convergence has practical usage for swapping integral for example. Moreover it gives two very important theorems: Arzela–Ascoli, Stone-Weierstrass.

Formal one is in the direction toward functional analysis. One can use Arzela–Ascoli in proving Picard–Lindelof theorem: existence solution in ode, where subsequence of sequence of functions converge to solution.

The later one is found usage in topological structure. Weierstrass used sequence of polynomials. The advantage over Taylor expansion is that doesn't require differentiability. But the price to pay is that the domain is bounded. Stone generalized that to a class of functions not necessary polynomials. This gives a lot of insight to Fourier series.

(2) Power Series, Fourier Series (chapter 8)

We will also discuss Laplace method (including Stirling formula). We will skip exponential and others in the chapter.

(3) Differentiability in many variable (chapter 9)

The chapter is not in the direction of functional analysis, getting more into differential geometry.

We will concentrate on

i) Inverse function theorem (the proof is pretty original, will take 2 lectures, and including all lemmas)

ii) Implicit function theorem (this becomes a corollary of i)) This theorem provides how to reduce variables. And it works for non linear functions. (if it is linear, one just use rank)

iii) Constant rank theorem. (this talks about dimension of manifold)

We will skip Lagrangian multiplier.

(4) Because the proof of Stokes theorem takes a lots time, we will only give statement of the theorem, then we'll have time to do measure theory (Fubini, product measure) and Lebesgue integral.

We'll follow Rudin II, because it was done much formally (category version) Rudin II than in Rudin I.

1 Sequences of functions

1.1 Uniform Convergence

Given a sequence of functions $f_n : E \to \mathbb{R}^k$ with (E, d) metric space and the codomain is complete. Pick a limit point of E, we can do 2 limits

$$\lim_{n\to\infty} \lim_{x\to p} f_n(x)$$

or

$$\lim_{x\to p} \lim_{n\to\infty} f_n(x)$$

As we will see, in general they do not give the same value.

Notice if $\lim_{n\to\infty} f_n(x)$ exists for all x, we can of course define

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

the pointwise limit.

We can present this issue in different flavors.

Consider

$$\lim_{n\to\infty} \int_a^b f_n(x)dx \quad \& \quad \int_a^b \lim_{n\to\infty} f_n(x)dx$$

The two don't agree, and it can be even worse that f_n integrable, but $f(x)$ is not.

Also

$$\lim_{n\to\infty} f_n'(x) \qquad \left(\lim_{n\to\infty} f_n(x)\right)'$$

f_n differentiable, but $f(x)$ is not.

This problem of swapping is beyond the boundary of analysis, in algebra, mostly swapping is not allowed too.

Example 1. Just some algebraic equation

$$\frac{n}{n+m}$$

doing limit $n \to \infty$ then take $\lim_{m\to\infty}$ is not the same as doing m first.

Example 2. More sophistical example let

$$f_n(x) = \sum_{j=0}^{n} \frac{x^2}{(1+x^2)^j}$$

$x \in \mathbb{R}$. We see that f_n is continuous for all $n \in \mathbb{N}$ and all x, but

$$f(x) = \lim_{n\to\infty} f_n = \begin{cases} 0 & x = 0 \\ x^2 \frac{1}{1-\frac{1}{1+x^2}} = 1 + x^2 & x \neq 0 \end{cases}$$

so not continuous at $x = 0$. Hence

$$\lim_{n\to\infty} \lim_{x\to 0} f_n(x) \neq \lim_{x\to 0} \lim_{n\to\infty} f_n(x)$$

Example 3. (Dirichlet function) Consider

$$f_m(x) = \lim_{n\to\infty} (\cos m!\pi x)^{2n}$$

$x \in [0, 1]$. It is easy to see if $x \notin \mathbb{Q}$,

$$|\cos m!\pi x| < 1$$

so $f_m = 0$. If $x = p/q \in \mathbb{Q}$, then for fix m, if $m!$ is divisible by q, $f_m = 1$. In other words

$$f_m = 1$$

at finitely many points of x, and $f_m = 0$ elsewhere. So f_m is Riemann integrable. But

$$\lim_{m\to\infty} f_m = \begin{cases} 0 & x \notin \mathbb{Q} \\ 1 & x \in \mathbb{Q} \end{cases}$$

This is Dirichlet function, not integrable.

This example is very hard to come up with, because not many commonly seen functions that are integrable but the limit is not integrable. Next example is not to hard to find.

Example 4. Let

$$f_n(x) = \frac{\sin nx}{\sqrt{n}}$$

then

$$f_n' = \sqrt{n} \cos nx$$

so f_n' is differentiable. But

$$\lim_{n\to\infty} f_n'$$

is not convergent for all x, hence

$$\lim_{n\to\infty} f_n'(x) \neq \left(\lim_{n\to\infty} f_n(x)\right)'$$

Example 5. Now we'll see some pretty common functions can do something very bad in terms of swapping. Let

$$f_n(x) = nx(1-x^2)^n$$

$x \in [0,1]$. This is of course familiar polynomials.

One will find

$$\int_0^1 \lim_{n\to\infty} f_n(x)dx = \int 0dx = 0$$

but

$$\lim_{n\to\infty} \int_0^1 f_n(x)dx = \lim_{n\to\infty} \frac{n}{2(n+1)} = \frac{1}{2}$$

notice everything here is integrable and everything is convergent, but still they don't agree.

Definition 6. $f_n : E \to \mathbb{R}^k$. f_n *converges pointwise* in $\mathbb{R}^k$ to f iff $\forall\, \varepsilon > 0$, $\forall\, x \in E$, $\exists\, N_{\varepsilon,x} \in \mathbb{N}$ s.t. $\forall\, n \geq N_{\varepsilon,x}$

$$|f_n(x) - f(x)| < \varepsilon$$

Definition 7. f_n *converges uniformly* to f iff $\forall\, \varepsilon > 0$, $\exists\, N_\varepsilon \in \mathbb{N}$ s.t. $\forall\, n \geq N_\varepsilon$

$$|f_n(x) - f(x)| < \varepsilon$$

$\forall\, x \in E$.

Remark 8. Uniformly convergence $\implies$ pointwise convergence. The converse is not true. See Examples 1-4.

Proposition 9. *$f_n \to f$ uniformly iff $\forall\, \varepsilon > 0$, $\exists\, N_\varepsilon \in \mathbb{N}$ s.t. $\forall\, n, m \in N_\varepsilon$*

$$|f_n(x) - f_m(x)| < \varepsilon \quad \forall\, x \in E$$

This says uniform convergence is the same as uniform Cauchy.

Proof. If $f_n(x) \to f(x)$ uniformly

$$|f_n(x) - f_m(x)| \leq |f_n - f| + |f - f_m| \leq \varepsilon$$

$\forall\, x \in E$, $\forall\, n, m \geq N_{\varepsilon/2}$.

Conversely if $|f_n - f_m| < \varepsilon \; \forall\, x \in E$. Fix x, $f_n(x)$ is Cauchy, so convergent ($\because$ space is complete), i.e. $\exists\, f(x)$ s.t.

$$f_n(x) \to f(x)$$

for this x. This defines $f(x) \; \forall\, x$. Then since

$$|f_n - f_m| < \varepsilon \quad \forall\, x \in E$$

$\forall\, m, n \geq N_\varepsilon$, take $m \to \infty$,

$$|f_n - f| < \varepsilon$$

the proposition is proved. ∎

Lecture 2 (1/24/13)

Remark 10. $f_n \to f$ uniformly iff

$$M_n = \sup_{x \in E} |f_n(x) - f(x)| \to 0 \text{ as } n \to \infty$$

Proposition 11. *$\sum_{n=0}^{\infty} f_n(x)$ and $\forall\, n \in \mathbb{N} \; \exists\, M_n$ s.t. $|f_n(x)| \leq M_n$ $\forall\, x \in E$ and $\sum_{n=0}^{\infty} M_n < \infty$, then $\sum_{n=0}^{\infty} f_n$ uniform convergent.*

Proof. Let q, $p > n$

$$\left| \sum_{k=p}^{q} f_k(x) \right| \leq \sum_{k=p}^{q} |f_n(x)| \leq \sum_{k=p}^{q} M_k < \varepsilon$$

$\forall\, x$, so $\sum f_n$ uniform Cauchy, hence uniform convergent. ∎

Theorem 12. *$f_n \to f$ uniformly for $x \in E$, p limit point of E. Assume that $\lim_{x \to p} f_n(x) = A_n < \infty \; \forall\, n$, then*

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

Proof. First show A_n convergent

$$|f_n(x) - f_m(x)| < \varepsilon$$

$\forall\, x \in E$ and $n, m > N_\varepsilon$. Taking limit $x \to p$ gives

$$|A_n - A_m| \leq \varepsilon$$

(notice usually limit changes $<$ to $\leq$) It follows $A_n \to A$.

We are done with the right hand of the conclusion, now prove that $\lim_{x\to p} f(x) = A$.

$$|f(x) - A| \leq |f(x) - f_n(x)| + |f_n(x) - A_n| + |A_n - A| \leq 3\varepsilon$$

The first and third terms in the middle are $\leq$ for $n \geq N_\varepsilon$ and $n \geq M_\varepsilon$ respectively. Then one can fix $n \geq \max\{N_\varepsilon, M_\varepsilon\}$. For this n, one finds δ_ε s.t.

$$d(x,p) < \delta_\varepsilon \implies |f_n(x) - A_n| < \varepsilon$$

□

Corollary 13. *$f_n \to f$ uniform, f_n continuous at p, then f is continuous at p.*

Proof.

$$\begin{aligned} \lim_{x\to p} f(x) &= \lim_{n\to\infty}\lim_{x\to p} f_n(x) \text{ by thm 12} \\ &= \lim_{n\to\infty} f_n(p) \text{ by hypothesis} \\ &= f(p) \text{ by definition of } f \end{aligned}$$

□

The following proposition will help us to check if something is uniform convergent without much calculation.

Proposition 14. *$f_n : K \to \mathbb{R}$, K compact. Assume*

(1) f_n are continuous (2) $f_n \to f$ pointwise and f is continuous (3) $f_n(x) \geq f_{n+1}(x)\ \forall\, x$, then $f_n \to f$ uniform.

Proof. let $g_n = f_n(x) - f(x)$, so g_n continuous, $g_n \geq g_{n+1}$ and $g_n \to 0$ pointwise $\forall\, x \in K$.

Fix $\varepsilon > 0$, let $K_n = g_n^{-1}[\varepsilon, \infty)$, then K_n is compact. Indeed, $[\varepsilon,\infty)$ is closed, and g_n continuous so K_n is closed, and $K_n \subset K$ closed in compact is compact.

Since $g_n \downarrow \implies K_n \supseteq K_{n+1}$. let $\tilde{K} = \bigcap_{n=1}^{\infty} K_n$, fix $x \in K$, $g_n(x) \to 0 \implies$

$$\exists\, N \text{ s.t. } g_n(x) < \varepsilon\ \forall\, n \geq N \implies x \notin K_n \text{ if } n \geq N$$

hence $x \notin \tilde{K}$, so $\tilde{K} = \varnothing \implies \exists\, M$ s.t. $K_n = \varnothing\ \forall\, n \geq M \implies 0 \leq g_n(x) \leq \varepsilon\ \forall\, n \geq M\ \forall\, x \in K$, that follows $g_n \to 0$ uniformly. □

The following example shows that K compact is very important.

Example 15. $f_n(x) = \frac{1}{1+nx}$ $x \in (0,1)$, we have $f_n \geq f_{n+1}$, $f_n \to 0$. But $\sup_{x\in(0,1)} f_n(x) = 1$, so not convergence uniform.

Theorem 16. *$f_n(x) \in \mathcal{R}(\alpha)$, Riemann integrable, on $[a,b]$ and $f_n(x) \to f(x)$ uniform, then*

(1) $f \in \mathcal{R}$ (2) $\int_a^b f d\alpha = \lim_{n\to\infty} \int_a^b f_n d\alpha$.

Proof. $\varepsilon_n = \sup_{x\in[a,b]} |f(x) - f_n(x)| \to 0$ as $n \to \infty$. That is

$$f_n(x) - \varepsilon_n \leq f(x) \leq f_n(x) + \varepsilon_n \quad \forall\, x$$

That implies

$$\int_a^b (f_n - \varepsilon_n) d\alpha \leq \underline{\int_a^b} f d\alpha \leq \overline{\int_a^b} f d\alpha \leq \int_a^b (f_n + \varepsilon_n) d\alpha \tag{1.1}$$

(recall although we don't know if f is integrable, we can take limsup/liminf), subtract the two outer integrals and use linearity of integration, we get

$$0 \leq \overline{\int_a^b} f d\alpha - \underline{\int_a^b} f d\alpha \leq 2\varepsilon_n(\alpha(b) - \alpha(a)) < \varepsilon$$

if $n \geq N_\varepsilon$. So we showed f is integrable.

Then subtracting $\int_a^b f_n$ from (1.1), to get

$$-\varepsilon_n(\alpha(b) - \alpha(a)) \leq \int_a^b (f - f_n) d\alpha \leq \varepsilon_n(\alpha(b) - \alpha(a))$$

Both outer quantities go to 0. □

Remark 17. Later in the semester, we will show that in the above theorem, f_n convergence uniform is not necessary, when we deal with Lebesgue.

The following swapping derivative theorem doesn't require f_n convergence uniform.

Theorem 18. *$f_n(x)$ differentiable on $[a,b]$ and $f_n'(x)$ convergence uniformly $\forall\, x$, and $\exists\, x_0 \in [a,b]$ s.t. $f_n(x_0)$ is convergent, then f_n converges uniformly to $f(x)$, and $f(x)$ is differentiable, and $f' = \lim_{n\to\infty} f_n'$.*

Proof. We show first f_n converges uniformly. This is the hardest part in the proof. We are given

$$\left|f_n'(x) - f_m'(x)\right| < \varepsilon$$

$$|f_n(x_0) - f_m(x_0)| < \varepsilon$$

$\forall\, x \in [a,b]\; m, n \geq N_\varepsilon$.

We use mean value theorem, (not on f_n, but $f_n - f_m$)

$$(f_n(t) - f_m(t)) - (f_n(x) - f_m(x)) = (f_n - f_m)'(y)(t - x)$$

It follows

$$\begin{array}{rcl} |(f_n(t) - f_m(t)) - (f_n(x) - f_m(x))| & = & \left|(f_n - f_m)'(y)\right| |t - x| (1.2) \\ & \leq & \varepsilon(b - a) \end{array}$$

$\forall\, t, x \in [a,b]$.

In particular if $x = x_0$, then

$$\begin{array}{rcl} |f_n(t) - f_m(t)| & \leq & |(f_n(t) - f_m(t)) - (f_n(x_0) - f_m(x_0))| + |f_n(x_0) - f_m(x_0)| \\ & \leq & \varepsilon(b - a) + \varepsilon \end{array}$$

So f_n converges uniformly.

Now fix $x \in [a,b]$, let $\psi_n(t) = \frac{f_n(t) - f_n(x)}{t - x}$ and then $\lim_{t \to x} \psi_n(t) = f_n'(x)$, use (1.2) again

$$|\psi_n(t) - \psi_m(t)| = \left|(f_n - f_m)'(y)\right| < \varepsilon\; \forall\, t$$

Since we know

$$\lim_{n \to \infty} \lim_{t \to x} \psi_n(t) = \lim_{n \to \infty} f_n'(x) = f'(x)$$

Applying theorem 12, $\lim_{t \to x} \lim_{n \to \infty} \psi_n(t)$ exists and it follows

$$\lim_{t \to x} \lim_{n \to \infty} f_n'(t) = \lim_{t \to x} \lim_{n \to \infty} \psi_n(t) = \lim_{n \to \infty} \lim_{t \to x} \psi_n(t) = f'(x)$$

$\square$

Corollary 19. $\sum_{n=0}^{\infty} f_n(x)$ *uniformly converges, then*

$$\lim_{x \to p} \sum_{n=0}^{\infty} f_n(x) = \sum_{n=0}^{\infty} \lim_{x \to p} f_n(x)$$

1.2 Arzela-Ascoli

Definition 20. $f : E \to \mathbb{R}^K$, E metric space, f continuous and bounded functions, i.e codomain is bounded in a ball.

$$\mathcal{C}(E) = \text{set of these functions}$$

$$d(f,g) = \sup_{x \in E} |f(x) - g(x)| \equiv \|f - g\|_\infty$$

One can check $\| \|_\infty$ is a distance, $(\mathcal{C}, \| \|_\infty)$ is a metric space.

Proposition 21. *$f_n \to f$ in $\| \|_\infty \iff f_n \to f$ uniformly.*

Proof. Let $f_n \in \mathcal{C}(E)$, $f_n \to f$ in $\| \|_\infty$

$$\sup |f_n(x) - f(x)| = \|f_n - f\|_\infty < \varepsilon$$

if $n > N_\varepsilon$. □

Lecture 3 (1/29/13)

Proposition 22. *$(\mathcal{C}(E), \| \|_\infty)$ is a complete metric space.*

Proof. If $f_n(x) \in \mathcal{C}(E)$ that converges uniformly to $f(x)$, $f(x)$ is continuous by corollary 13.

f_n is bounded implies $f(x)$ is bounded. Indeed we can find N s.t.

$$|f_n(x) - f(x)| < 1$$

$\forall\, x \in E$, $\forall\, n \geq N$. □

We now have a nice metric space to work with. Recall a bounded sequence in E has a convergent subsequence. E is not required to be compact. What can we say about sequence of functions?

Definition 23. Bounded sequence $f_n(x)$ is *pointwise bounded* iff $\forall\, x \in E$, $\exists\, M_x \in \mathbb{R}$ s.t. $|f_n(x)| \leq M_x$ $\forall\, n \in \mathbb{N}$.

Definition 24. $f_n(x)$ is *uniformly bounded* if $\exists\, M \in \mathbb{R}$ s.t. $|f_n(x)| \leq M$ $\forall\, x, n$.

Remark 25. Notice that even a uniformly bounded $f_n(x) \not\Longrightarrow \exists$ convergent subsequence.

e.g.

$$f_n(x) = \sin nx$$

$-\pi \leq x \leq \pi$, f_n is uniformly bounded. Assume $\exists\, n_k$ s.t.

$$f_{n_k} = \sin n_k x$$

is convergent, then we could have

$$(\sin(n_{k+1}x) - \sin(n_k x))^2 \to 0 \tag{1.3}$$

But

$$\lim_{k\to\infty} \int_{-\pi}^{\pi} (\sin(n_{k+1}x) - \sin(n_k x))^2 dx = 2\pi$$

because the integral is independent of n_k, while as

$$\int_{-\pi}^{\pi} \lim_{k\to\infty} (\sin(n_{k+1}x) - \sin(n_k x))^2 dx = 0$$

the swapping of limit and integral is given by Lebesgue dominated convergence theorem, which we will study near the end of the semester.

This example shows that even pointwise convergent subsequence is not guaranteed.

Now let's tackle this problem. We use a notion of *family* of functions, slightly more general than sequence of functions.

Definition 26. $\mathcal{F}$ a family of functions $f : E \to \mathbb{R}^K$. $\mathcal{F}$ is *equicontinuous* iff $\forall\, \varepsilon > 0$, $\exists\, \delta_\varepsilon > 0$ s.t.

$$d(x, y) < \delta_\varepsilon \implies |f(x) - f(y)| < \varepsilon$$

$\forall\, f \in \mathcal{F}$.

Proposition 27. *K is compact, $f_n(x)$ continuous functions, and $f_n(x)$ converges uniformly, then $\{f_n(x)\}_{n\in\mathbb{N}}$ is equicontinuous.*

Proof. Fix $\varepsilon > 0$

$$|f_n(x) - f_m(x)| < \varepsilon$$

$\forall\, x \in K$, $\forall\, n, m \geq N_\varepsilon$.

Want to show $\exists\, \delta > 0$ s.t. if $d(x, y) < \delta$,

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

$\forall\, n \in \mathbb{N}$.

If $i \leq N_\varepsilon$, we can find δ_i s.t.

$$d(x,y) < \delta_i \implies |f_i(x) - f_i(y)| < \varepsilon$$

because f_i is continuous on K, i.e. f_i is uniformly continuous. Let

$$\delta = \min\{\delta_1, ..., \delta_{N_\varepsilon}\}$$

so if $n \leq N_\varepsilon$,

$$d(x,y) < \delta \implies |f_n(x) - f_n(y)| < \varepsilon$$

If $n > N_\varepsilon$, $d(x,y) < \delta$, then

$$\begin{aligned} |f_n(x) - f_n(y)| &\leq |f_n(x) - f_{N_\varepsilon}(x)| + |f_{N_\varepsilon}(x) - f_{N_\varepsilon}(y)| + |f_{N_\varepsilon}(y) - f_n(y)| \\ &\leq 3\varepsilon \end{aligned}$$

The 1st and 3rd term on the right is due to uniform convergent of f_n, and the middle term is due to continuity of f_n □

Theorem 28. *(Arzela–Ascoli) K compact, $f_n(x)$ a sequence of functions on K, which are equicontinuous and pointwise bounded, then*

(1) $f_n(x)$ is uniformly bounded (2) $\exists$ a subsequence of functions that converges uniformly.

Lemma 29. *$f_n(x)$ a sequence of functions pointwise bounded on a countable domain E, then $\exists$ a pointwise convergent subsequence.*

This lemma is not so useful in practice, but we will just the lemma to build a candidate for the convergent subsequence. Recall a compact set is separable by countable balls.

Proof. (of lemma) $\{x_n\}_{n\in\mathbb{N}} = E$, $|f_n(x_i)| \leq M_i \; \forall x_i \in E$, then there exists a subsequence of f_n, denoted by

$$f_{11}(x),\, f_{12}(x),\, ...,\, f_{1n}(x), ...$$

converges on x_1, then take a subsequence of the above subsequence, denoted by

$$f_{21}(x),\, f_{22}(x),\, ...,\, f_{2n}(x), ...$$

converges on $x_{2,}$ continue this process, let

$$g_n(x) = f_{nn}(x)$$

then g_n converges $\forall\, x$.

Proof. (of Arzela-Ascoli) (1) f_n pointwise bounded and equicontinuous, i.e.

$$|f_n(x)| \leq M_x$$

$\forall\, n$, and $\forall\, \varepsilon > 0\ \exists\, \delta > 0$ s.t.

$$d(x,y) < \delta \implies |f_n(x) - f_n(y)| < \varepsilon$$

Want to show

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

$\forall\, n, x$.

Given a cover of K

$$K \subset \bigcup_{x \in K} B_{x,\delta}$$

then one finds a finite subcover

$$K \subset \bigcup_{i=1}^{m} B_{x_i,\delta}$$

Let

$$M = \max\{M_{x_1}, ..., M_{x_m}\}$$

For $x \in B_{x_i,\delta}$

$$|f_n(x)| \leq |f_n(x) - f_n(x_i)| + |f_n(x_i)|$$

where

$$|f_n(x) - f_n(x_i)| < \varepsilon$$

if $d(x, x_i) < \delta$

$$|f_n(x_i)| \leq M_{x_i} \leq M$$

so

$$|f_n(x)| \leq \varepsilon + M$$

$\forall\, n, x$.

(2) K compact implies K separable, so we can find

$$E \subset K;\ E \text{ is dense and countable}$$

then by the lemma, $g_n(x)$ is a subsequence of $f_n(x)$ that converges $\forall\, x \in E$.

Want to show

$$|g_n(x) - g_m(x)| < \varepsilon$$

$\forall\, x \in K$, $m, n \geq N_\varepsilon$.

We have

$$|g_n(x_i) - g_m(x_i)| < \varepsilon$$

$x_i \in E$, $m, n \geq N_{\varepsilon,i}$.

Fix ε, $\exists\, \delta > 0$

$$K \subset \bigcup_{x \in E} B_{x,\delta}$$

covers K, because E is dense in K. Take finite subcover

$$K \subset \bigcup_{i=1}^{m} B_{x_i,\delta}$$

$N_\varepsilon \equiv \max\{N_{\varepsilon 1}, N_{\varepsilon 2}, ..., N_{\varepsilon m}\}$, then for $x \in B_{x_i,\delta}$,

$$\begin{aligned} |g_n(x) - g_m(x)| &\leq |g_n(x) - g_n(x_i)| + |g_n(x_i) - g_m(x_i)| + |g_m(x_i) - g_m(x)| \\ &\leq 3\varepsilon \end{aligned}$$

the 1st and 3rd terms on the right are due to equicontinuity, and the middle term is given by the lemma when $m, n \geq N_\varepsilon$. □

□

For application of Arzela-Ascoli, prove Picard-Lindelof theorem, read the last two problems outlined in Rudin chapter 7.

1.3 Stone-Weierstrass

Lecture 4 (1/29/13)

We will continuous work in the metric space $(\mathcal{C}[a,b], \|\ \|_\infty)$. Now we discuss another important property of this metric space.

Theorem 30. *(Weierstrass) $f \in \mathcal{C}[a,b]$ is real-valued. $\mathcal{P}$ set of all polynomials on $[a,b]$, then*

$$\mathcal{P} \textit{ is dense in } \mathcal{C}[a,b]$$

or equivalently

$$\bar{\mathcal{P}} = \mathcal{C}[a,b]$$

or equivalently $\forall\, f \in \mathcal{C}[a,b]\ \exists\, P_n \in \mathcal{P}$ a sequence of polynomials s.t.

$$P_n \to f$$

in $\|\ \|_\infty$, i.e. uniformly.

Later we will generalize this to Stone, where we will allow complex-valued functions. The proof given here is followed from Rudin. There are other proofs that are more direct, and use piecewise linear function (polygon curves).

Proof. First reduce the problem to $[0, 1]$

$$\begin{aligned} [0,1] &\rightarrow [a,b] \\ t &\mapsto (1-t)a + bt \end{aligned}$$

notice it is a bijection and a polynomial (linear) transformation.

Second we can assume $f(a) = f(b) = 0$. If not, we can define

$$g(t) = (b-a)(f(t) - f(a)) - (t-a)(f(b) - f(a))$$

notice g is a polynomial.

Now $f(t)$ on $[0, 1]$ and $f(0) = f(1) = 0$, extend f

$$\tilde{f}(t) = \begin{cases} f & t \in [0,1] \\ 0 & t \notin [0,1] \end{cases}$$

$\tilde{f} \in \mathcal{C}(\mathbb{R})$. Define

$$Q_n = c_n(1-t^2)^n$$

where $c_n \in \mathbb{R}$ s.t.

$$\int_{-1}^{1} Q_n dt = 1$$

Let

$$P_n(x) = \int_{-1}^{1} Q_n(t) f(t+x) dt$$

$x \in [0, 1]$, $n \in \mathbb{N}$

(Normally it is called "Convolution". It is just projection of $f(x)$ onto orthonormal basis Q_n. Similar idea $c_n e^{inx}$,$n \in \mathbb{Z}$ Fourier basis.)

Let's find an estimate for c_n. (We don't need the actual expression, too complicated)

$$1 = c_n \int_{-1}^{1} (1-t^2)^n dt = 2c_n \int_{0}^{1} (1-t^2)^n dt \geq 2c_n \int_{0}^{1/\sqrt{n}} (1-nt^2) dt$$

recall assume $a \geq b \geq 0$,

$$\begin{aligned} a^n - b^n &\leq (a-b)(a^{n-1} + ba^{n-2} + ...) \\ &\leq (a-b)na^{n-1} \end{aligned}$$

So here

$$1^n - (1-t^2)^n \leq t^2 n$$

or

$$(1-t^2)^n \geq 1 - nt^2$$

We also changed the limit of the integration from 0 to $1/\sqrt{n}$, to make the integrand stays positive. So we get

$$1 \geq c_n \frac{3}{4\sqrt{n}} > c_n \frac{1}{\sqrt{n}} \implies c_n < \sqrt{n}$$

Prove $P_n(x)$ is a polynomial, $x \in [0,1]$

$$\begin{aligned} P_n(x) &= \int_{-1}^{1} Q_n(t) f(t+x) dt \text{ set } t+x=s \\ &= \int_{-1+x}^{1+x} Q_n(s-x) f(s) ds \because f(s) = 0 \text{ for } s \notin [0,1] \\ &= \int_0^1 Q_n(s-x) f(s) ds \end{aligned}$$

which is clearly a polynomial.

Prove uniform convergence

Fix $\varepsilon > 0$,

$$|P_n(x) - f(x)| = \left| \int_{-1}^{1} f(t+x) Q_n(t) dt - \int_{-1}^{1} f(x) Q_n(t) dt \right|$$

(this trick will be used many times in our course, particular when we show the convergence of Fourier series)

Find $\delta > 0$ s.t.

$$|f(u) - f(v)| < \varepsilon \tag{1.4}$$

$\forall\, u, v$ s.t.$|u - v| < \delta$.

Therefore we can break

$$\begin{aligned} |P_n(x) - f(x)| \leq& \int_{-1}^{\delta} |f(t+x) - f(x)|\, Q_n(t) dt + \int_{-\delta}^{\delta} |f(t+x) - f(x)|\, Q_n(t) dt \\ &+ \int_{\delta}^{1} |f(t+x) - f(x)|\, Q_n(t) dt \end{aligned}$$

For $t \in [\delta, 1]$

$$Q_n(t) = c_n(1-t^2)^n \le \sqrt{n}(1-t^2)^n \le \sqrt{n}(1-\delta^2)^n \to 0$$

converges to 0 uniformly (i.e. independent of t), because $(1-\delta^2) < 1$, $(1-\delta^2)^n$ decays much faster than the growth of $\sqrt{n}$.

$M \equiv \sup_{x\in[0,1]} |f|$, so

$$\int_\delta^1 |f(t+x) - f(x)|\, Q_n(t)dt \le 2M \int_\delta^1 Q_n(t)dt < 2M\varepsilon$$

Similarly for $x \in [-1, -\delta]$ integral.

By (1.4)

$$\int_{-\delta}^{\delta} |f(t+x) - f(x)|\, Q_n(t)dt < \varepsilon \int_{-\delta}^{\delta} Q_n(t)dt < \varepsilon$$

hence

$$|P_n(x) - f(x)| < 4M\varepsilon + \varepsilon$$

$\square$

Stone generalized the result to any compact domain. e.g. a circle. This is a Fourier domain. Later we will show $f(x)$ is continuous and periodic can be written as a continuous function on S^1.

Definition 31. $\mathcal{A}$ *commutative algebra* is a vector space with a product

$$\mathcal{A} \times \mathcal{A} \to \mathcal{A}$$

$$a, b \mapsto a \cdot b$$

the product is associative, commutative, distributive and has a neutral element.

Theorem 32. *(Stone) K compact, metric space $(\mathcal{C}(K), \|\ \|_\infty)$, if real valued functions $\mathcal{A} \subset \mathcal{C}(K)$ an algebra*

(1) separates points (2) doesn't vanish, then

$$\bar{\mathcal{A}} = \mathcal{C}(K)$$

Just like in Weierstrass, $\bar{\mathcal{A}}$ means uniform closure, i.e. $f \in \bar{\mathcal{A}} \iff \exists\, g \in \mathcal{A},\ g \to f$ uniformly.

Definition 33. $\mathcal{A}$ separates points iff $\forall\, x_1 \neq x_2 \in K$, $\exists\, f \in \mathcal{A}$ s.t.

$$f(x_1) \neq f(x_2)$$

Definition 34. $\mathcal{A}$ doesn't vanish iff $\forall\, x \in K\ \exists\, f \in \mathcal{A}$ s.t.

$$f(x) \neq 0$$

The proof of Stone is less random then the proof of Weierstrass, and the proof is very enlightening. Before we do that, we need 3 lemmas.

Lemma 35. *If $\mathcal{A}$ is an algebra, then $\bar{\mathcal{A}}$ is also an algebra.*

Proof. We want to show $f, g \in \bar{\mathcal{A}}$,

$$\alpha f + \beta g \in \bar{\mathcal{A}}$$

$$fg \in \bar{\mathcal{A}}$$

There are $f_n, g_n \in \mathcal{A}$ s.t.

$$f_n \to f \quad g_n \to g$$

uniformly. So by triangle inequality,

$$\sup |(\alpha f_n + \beta g_n) - (\alpha f + \beta g)| \to 0$$

Similarly showing

$$\sup |f_n g_n - fg| \to 0$$

□

Lemma 36. *If $\mathcal{A}$ separates points and doesn't vanish, then $\forall\, c_1, c_2 \in \mathbb{R}$ or $\in \mathbb{C}$ and $\forall\, x_1, x_2 \in K$, $x_1 \neq x_2$, $\exists\, f \in \mathcal{A}$ s.t.*

$$f(x_1) = c_1 \qquad f(x_2) = c_2$$

Recall polynomial has this property, called polynomial interpolation. So set of all polynomials is an example of such $\mathcal{A}$.

Proof. $\exists\, g, h, k \in \mathcal{A}$

$$g(x_1) \neq g(x_2) \quad h(x_1) \neq 0 \quad k(x_2) \neq 0$$

Let

$$f_1(x) = h(x)g(x) - g(x_2)h(x)$$

$$f_2(x) = k(x)g(x) - g(x_1)k(x)$$

then $f_1(x_1) \neq 0$, $f_2(x_2) \neq 0$, and $f_1(x_2) = f_2(x_1) = 0$, so

$$f(x) = c_1 \frac{f_1(x)}{f_1(x_1)} + c_2 \frac{f_2(x)}{f_2(x_2)}$$

is the desired function. $\square$

Lemma 37. *Consider $|x|$ on $[a, b]$. $\exists$ a sequence of polynomials $P_n(x)$ s.t.*

$$P_n(x) \to |x|$$

as $n \to \infty$, and $P_n(0) = 0 \; \forall n \in \mathbb{N}$.

Proof. By Weierstrass, there are R_n, sequence of polynomials,

$$R_n \to |x|$$

uniformly. Let

$$P_n(x) = R_n(x) - R_n(0)$$

since $R_n(0) \to 0$, so P_n does the job. $\square$

We now prove Stone.

Proof. First show

$$f \in \bar{\mathcal{A}} \implies |f| \in \bar{\mathcal{A}}$$

Let $\alpha = \sup |f(x)|$, fix $\varepsilon > 0$, apply lemma 37 to find $P_n(x)$ for all $x \in [-\alpha, \alpha]$ s.t.

$$|P_n(x) - |x|| < \varepsilon$$

then substitute f for x, get

$$\left| \sum_{k=1}^{n} c_k f^k(x) - |f| \right| < \varepsilon$$

$\sum_{k=1}^{n} c_k f^k(x) \in \bar{\mathcal{A}}$, so $|f| \in \bar{\mathcal{A}}$. (notice we impose $P_n(0) = 0 \iff c_0 = 0$ in the polynomial is to take care of the case that c_0 may not be in $\bar{\mathcal{A}}$.)

Lecture 5 (2/5/13)

Next we show

$$f_1, f_2 \in \bar{\mathcal{A}} \implies \max\{f_1, f_2\}, \min\{f_1, f_2\} \in \bar{\mathcal{A}}$$

Indeed

$$\max\{f_1, f_2\} = \frac{f_1 + f_2}{2} + \frac{|f_1 - f_2|}{2} \in \bar{\mathcal{A}}$$
$$\min\{f_1, f_2\} = \frac{f_1 + f_2}{2} - \frac{|f_1 - f_2|}{2} \in \bar{\mathcal{A}}$$

Our goal is for given $f(x) \in C(K)$, fix $\varepsilon > 0$, can we find $g(x) \in \bar{\mathcal{A}}$, s.t.

$$|g(x) - f(x)| < \varepsilon \quad \forall\, x \in K$$

This implies $f(x)$ is a limit point of $\bar{\mathcal{A}}$, so $f(x) \in \bar{\mathcal{A}}$. Hence $C(K) \subset \bar{\mathcal{A}}$.

Since by corollary 13

$$\mathcal{A} \subset \bar{\mathcal{A}} \subset C(K)$$

Therefore $\bar{\mathcal{A}} = C(K)$.

Fix $x \in K$, for each $t \in K$, find $g_{x,t}(y) \in \bar{\mathcal{A}}$ s.t.

$$\begin{cases} g_{x,t}(x) = f(x) \\ g_{x,t}(t) > f(t) - \varepsilon \end{cases}$$

That is possible by lemmas 35, 36. Because f, $g_{x,t}$ are continuous, $\exists\, B_t$ a neighborhood of t s.t.

$$g_{x,t}(y) > f(y) - \varepsilon$$

$\forall\, y \in B_t$.

Since $\bigcup_{t\in K} B_t$ is an open cover, then there exists subcover $\bigcup_i^n B_{t_i} \supset K$, we want $g_x(y) \in \bar{\mathcal{A}}$ s.t.

$$\begin{cases} g_x(x) = f(x) \\ g_x(t) > f(t) - \varepsilon \quad \forall\, t \in K \end{cases}$$

take

$$g_x = \max\{g_{x,t_1}, ..., g_{x,t_n}\}$$

then

$$g_x(y) \geq g_{x,t_i}(y) > f(y) - \varepsilon\; \forall\, i = 1, ..., n$$

and

$$g_x \in \bar{\mathcal{A}}$$

we want $g \in \bar{\mathcal{A}}$ s.t.

$$f(t) + \varepsilon > g(t) > f(t) - \varepsilon$$

Because $g_x(x) = f(x) < f(x) + \varepsilon$, since g_x, f are continuous, $\exists\, B_x$, a neighborhood of x s.t. $\forall\, t \in B_x$

$$g_x(t) < f(t) + \varepsilon$$

And since $\bigcup_{x \in K} B_x$ is an open cover, there exists $\bigcup_{i=1}^{m} B_{x_i} \subset K$, let

$$g = \min\{g_{x_1}, ..., g_{x_m}\} \in \bar{\mathcal{A}}$$

we get

$$f(t) - \varepsilon < g(t) < g_{x_i}(t) < f(t) + \varepsilon \quad \forall\, i$$

So f is a limit point of $\bar{\mathcal{A}} \implies \bar{\mathcal{A}} = C(K)$. □

Theorem 38. *(Stone, complex version) K compact, $C(K)$ continuous complex valued function with $\| \; \|_\infty$. $\mathcal{A} \subset C(K)$ algebra and $\mathcal{A}$ separates points and doesn't vanish; if $f \in \mathcal{A} \implies \bar{f} \in \mathcal{A}$, then*

$$\bar{\mathcal{A}} = C(K)$$

Remark 39. The addition assumption: $\mathcal{A}$ is self-adjoint, i.e. $f \in \mathcal{A} \implies \bar{f} \in \mathcal{A}$ is important, we will see it is used in the proof. Notice that is the reason in complex Fourier series, one always sums from $-\infty$ to ∞ to include the conjugate. If one just does

$$\sum_{0}^{\infty}$$

not to include the conjugate, then the series is not convergent.

Proof. $f \in C(K)$ so $f = u + iv$, where $u, v \in C_{\mathbb{R}}(K)$ (continuous real valued functions)

If $f \in \mathcal{A}$, then

$$\frac{1}{2}(f + \bar{f}) = u \in \mathcal{A}$$

$$\frac{1}{2}(f - \bar{f}) = v \in \mathcal{A}$$

(that is why we require $\bar{f}$ in $\mathcal{A}$)

Let $\mathcal{A}_{\mathbb{R}}$ be the set of all real-valued functions on K which belong to $\mathcal{A}$.

Claim: $\mathcal{A}_{\mathbb{R}}$ separates points. $x_1 \neq x_2 \in K$ (by lemma 36) $\exists\, f \in \mathcal{A}$

$$f(x_1) = 1 \qquad f(x_2) = 0$$

and Re$f \in \mathcal{A} \implies$ Re$f \in \mathcal{A}_\mathbb{R} \implies \mathcal{A}_\mathbb{R}$ separates points.

Claim: $\mathcal{A}_\mathbb{R}$ doesn't vanish. $\forall\, x \in K$, $\exists\, f \in A$ s.t. $f(x) \neq 0$. Then

$$\overline{f(x)}f(x) = |f(x)|^2 > 0 \implies \text{Re}(f)(x) \neq 0$$

so $\mathcal{A}_\mathbb{R}$ doesn't vanish.

Use Stone on $\mathcal{A}_\mathbb{R}$, $\implies \bar{\mathcal{A}}_\mathbb{R} = C_\mathbb{R}(K)$. For $f \in C(K)$, $f = u + vi$,

$$u, v \in C_\mathbb{R}(K) \implies u, v \in \bar{\mathcal{A}}_\mathbb{R} \subset \bar{\mathcal{A}} \implies f \in \bar{\mathcal{A}}$$

□

2 Some Special Functions

2.1 Analytic Functions

Definition 40. Open $\Omega \subset \mathbb{R}$ or $\mathbb{C}$, $f(x)$ is analytic in Ω iff $\forall\, B_{a,r} \subset \Omega$

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n \quad \forall\, x \in B_{a,r}$$

Here Ω is not necessarily connected. Later we will show uniqueness of such representation and also show possibilities to extend the function outsider the ball.

Proposition 41. *$f(x) = \sum_{n=0}^{\infty} c_n x^n$ convergent if $|x| < R$, $\forall\, \varepsilon > 0$, $f(x)$ is uniformly convergent on $|x| \leq R - \varepsilon$.*

This is saying it is convergent in a closed ball that is contained in the open ball B_R.

Proof. By the hypotheses

$$\sum_{n=0}^{\infty} c_n (R-\varepsilon)^n < \infty \,\forall\, \varepsilon > 0$$

This is saying by the root test,

$$\lim_{n\to\infty} \sup \sqrt[n]{|c_n|}(R-\varepsilon) = \alpha(R-\varepsilon) < 1$$

so for $|x| \leq R - \varepsilon$

$$\lim_{n\to\infty} \sup \sqrt[n]{|c_n|}\,|x| = \alpha\,|x| < 1 \implies \sum |c_n|\,|x|^n \text{ convergent}$$

so it converges absolutely. □

Proposition 42. *$f(x) = \sum_{n=0}^{\infty} c_n x^n$ converges on $B_{0,R}$, $f(x)$ is differentiable on $B_{0,R}$ and its derivative is*

$$f'(x) = \sum_{n=1}^{\infty} c_n n x^{n-1}$$

Definition 43. (Complex derivative) $f : \Omega \to \mathbb{C}$, $\Omega \subset \mathbb{C}$ open $x_0 \in \Omega$

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

this is a limit in $\mathbb{R}^2$ Euclidean space.

The proposition works for real and complex, but here we will only give a proof for the real case. The proof for complex case is very different.

Proof. (in $\mathbb{R}$) Let $g(x) = \sum_{n=1}^{\infty} c_n n x^{n-1}$ in $B_{0,R}$

$$\lim_{n \to \infty} \sup \sqrt[n]{|c_n|} \sqrt[n]{n} = \lim_{n \to \infty} \sup \sqrt[n]{|c_n|}$$

so it has the same radius of convergence. Thus

$$\frac{d}{dx} \sum_{n=0}^{N} c_n x^n = \sum_{n=1}^{N} c_n n x^{n-1}$$

converges uniformly in $\overline{B_{0,R-\varepsilon}}$ $\forall \varepsilon > 0$, as $N \to \infty$. $\implies$ $f'(x) = \sum_{n=1}^{\infty} c_n n x^{n-1}$ in $\overline{B_{0,R-\varepsilon}}$ $\forall \varepsilon > 0$.

If $x \in B_{0,R}$, one can find a $\varepsilon > 0$ s.t. $x \in \overline{B_{0,R-\varepsilon}}$, so

$$f'(x) = \sum_{n=1}^{\infty} c_n n x^{n-1}$$

$\forall x \in B_{0,R}$. □

Remark 44. That the derivative of an analytic function is analytic implies f is analytic $\implies$ $f \in C^{\infty}$. The converse is true for complex function, but not true for real function in 1 variable, (note: this also reiterates the fact that real functions in 1 variable are not a subset of complex functions.) e.g.

Bump function

$$f = \begin{cases} 0 & x < -1 \text{ or } x > 1 \\ e^{-\frac{1}{1-x^2}} & -1 \le x \le 1 \end{cases}$$

this is a C^∞ function, but if one tries to find the Taylor series for it at center ± 1, one gets all terms including all derivatives to be 0.

This function doesn't satisfy a condition that we will discuss next lecture: if f is analytic $Z(f) = \{x \in \Omega \text{ s.t. } f(x) = 0\}$ cannot have limit point unless $f(x) \equiv 0$.

Notice proposition 42 also provides an important point about Taylor coefficients,

$$f(x) = \sum_{n=0}^{\infty} c_0 x^n \implies c_0 = f(0)$$

$$f'(x) = \sum_{n=1}^{\infty} c_n n x^{n-1} \implies c_1 = f'(0)$$

$$c_2 = \frac{f''(0)}{2!}$$

and so on.

Lecture 6 (2/7/13)

Corollary 45. *$f(x) = \sum_{n=0}^{\infty} c_0 x^n$ analytic on $|x| < R$, it can be differentiate arbitrary many times and*

$$c_n = \frac{f^{(n)}(0)}{n!}.$$

Today's task is to change the center of Taylor series to a point rather than 0. To do so, we need develop several propositions.

First consider $f(x) = \sum c_n x^n$ converges on $|x| < R$ (by the root test) and suppose

$$\sum c_n R^n = S$$

also converges (by some other tests) or

$$\sum c_n (-R)^n = T,$$

then is true

$$\lim_{x \to R^-} \text{ or } \lim_{x \to -R^+} f(x) = S \text{ or } T?$$

The answer is positive.

Theorem 46. *(Abel)* $\sum_{n=0}^{\infty} c_n < \infty$*;* $f(x) = \sum_{n=0}^{\infty} c_n x^n$ $|x| < 1$*, then*

$$\lim_{x \to 1^-} f(x) = \sum_{n=0}^{\infty} c_n$$

i.e. $f(x)$ *is continuous at 1.*

Remark 47. The hypothesis implies $f(x)$ is absolutely convergent. Proof: want to show

$$\lim_{n \to \infty} \sup \sqrt[n]{|c_n|}|x| < 1$$

Clearly $\sum_{n=0}^{\infty} c_n < \infty \implies \lim_{n\to\infty} \sup \sqrt[n]{|c_n|} \leq 1$, because it cannot be > 1.

Proof. Let $S_N = \sum_{n=0}^{N} c_n \to S$ for $N \geq 0$, and $S_{-1} = 0$, then (the same tricky used to prove conditional convergence, one term of the product is bounded, the other monotone decreases to 0.)

$$\begin{aligned}
\sum_{n=0}^{m} c_n x^n &= \sum_{n=0}^{m} (S_n - S_{n-1}) x^n \\
&= \sum_{n=0}^{m} S_n x^n - \sum_{n=0}^{m-1} S_n x^{n+1} \\
&= \sum_{n=0}^{m} S_n (x^n - x^{n+1}) + S_m x^{m+1} \\
&= (1 - x) \sum_{n=0}^{m} S_n x^n + S_m x^{m+1}
\end{aligned}$$

Because $S_m \to S$, i.e. bounded, $x^{m+1} \to 0$ ($\because |x| < 1$),

$$f(x) = (1 - x) \sum_{n=0}^{\infty} S_n x^n \quad |x| < 1$$

Since

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1 - x} \implies (1 - x) \sum_{n=0}^{\infty} S x^n = S$$

Consider

$$|f(x) - S| = |1 - x| \left| \sum_{n=0}^{\infty} (S_n - S) x^n \right|$$

Fix $\varepsilon > 0$, find N_ε s.t.

$$|S_n - S| < \varepsilon \quad \text{if } n > N_\varepsilon$$

Then we get

$$\begin{aligned} |f(x) - S| &\leq |1-x| \left| \sum_{n=0}^{N_\varepsilon} (S_n - S) x^n \right| + |1-x| \left| \sum_{N_\varepsilon+1}^{\infty} (S_n - S) x^n \right| \\ &\leq |1-x| \sum_{n=0}^{N_\varepsilon} |S_n - S| \, |x|^n + |1-x| \, \varepsilon \sum_{N_\varepsilon+1}^{\infty} |x|^n \end{aligned}$$

Without last of generality, ($\because$ we only interested in $x \to 1^-$) assume $0 < x < 1$, then

$$\sum_{N_\varepsilon+1}^{\infty} (1-x) x^n \leq 1$$

It follows

$$|f(x) - S| \leq (1-x) \sum_{n=0}^{N_\varepsilon} |S_n - S| + \varepsilon$$

Clearly

$$\sum_{n=0}^{N_\varepsilon} |S_n - S| < M_\varepsilon$$

so choose $\delta_\varepsilon > 0$ s.t.

$$(1-x) M_\varepsilon < \varepsilon \quad \text{if } 1 - \delta_\varepsilon < x < 1$$

$\square$

Recall product of two power series may not converge, but we have a theorem says if one of them is absolutely converges then the product converges. Use Abel, we get slightly different scope, although it is not more more general than what we have before.

Proposition 48. *$\sum_{n=0}^{\infty} a_n = A$, $\sum_{n=0}^{\infty} b_n = B$, let $c_n = \sum_{m=0}^{n} a^m b^{n-m}$ assume that $\sum_{n=0}^{\infty} c_n = C$, then*

$$AB = C$$

Proof. Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ by root test it's absolutely convergent if $|x| < 1$; let $g(x) = \sum_{n=0}^{\infty} b_n x^n$, $h(x) = \sum_{n=0}^{\infty} c_n x^n$ same radii of convergence. Hence the product of two absolutely convergent series is convergent, i.e.

$$f(x)g(x) = h(x) \text{ if } |x| < 1$$

Now use Abel, apply limit $x \to 1^-$ to both sides, (notice for Abel to work we need this additional condition $\sum_{n=0}^{\infty} c_n < \infty$), then we have the desired results. □

Another interesting proposition about power series is for a double summation, when we're allowed to swap two indices. We know from lecture one, example 1, swapping limit is not allowed in general for sequence. Since any sequence can be written as series vise versa, by telescopic way. We need the following "discrete version" of Fubini-Tonelli.

Proposition 49.

$$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_{ij} = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} a_{ij}$$

if $\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} |a_{ij}| < \infty$.

The proof is tricky; it uses sequence of functions, so we can take advantage of the criterion for uniform convergence.

Proof. Consider x_n, a sequence and $x_n \to x_0$, let

$$E = \{x_n\}_{n\in\mathbb{N}} \cup \{x_0\}$$

So E is compact, i.e. closed bounded, has only one limit point.

Let

$$f_i : E \to \mathbb{C} \text{ or } \mathbb{R}$$

with

$$f_i(x_n) := \sum_{j=0}^{n} a_{ij} \qquad n \geq 1$$

$$f_i(x_0) := \sum_{j=0}^{\infty} a_{ij}$$

hence

$$\lim_{n\to\infty} f_i(x_n) = f_i(x_0) \implies$$

f_i is continuous at x_0, (in fact it is continuous everywhere in E, $\because$ other points are isolated.)

Define

$$g(x) := \sum_{i=0}^{\infty} f_i(x) \quad x \in E$$

similarly, hence

$$g(x_0) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_{ij}$$

$$g(x_n) = \sum_{i=0}^{\infty} \sum_{j=0}^{n} a_{ij} = \sum_{j=0}^{n} \sum_{i=0}^{\infty} a_{ij}$$

Swap is allowed, because it's a finite sum.

Now we want to show $\sum_{i=0}^{\infty} f_i(x)$ is uniformly convergent (i.e. independent of x). Indeed

$$|f_i(x)| \leq \sum_{j=0}^{\infty} |a_{ij}| = b_i \text{ and } \sum b_i < \infty$$

That $\sum f$ uniformly convergent and $f_i(x)$ continuous at x_0 imply $g(x)$ is continuous at x_0. That is

$$\lim_{n\to\infty} g(x_n) \quad = \quad g(x_0)$$

and

$$\lim_{n\to\infty} g(x_n) \quad = \quad \lim_{n\to\infty} \sum_{j=0}^{n} \sum_{i=0}^{\infty} a_{ij}$$

$\square$

Proposition 50. *(Taylor) $f(x) = \sum_{n=0}^{\infty} c_n x^n$ converges on $|x| < R$, $x \in \mathbb{C}$ or $\mathbb{R}$, then we can always write*

$$f(x) = \sum_{n=0}^{\infty} d_n (x-a)^n$$

converges on $|x-a| < R - |a|$.

Proof. Use binomial

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} c_n (x-a+a)^n \\ &= \sum_{n=0}^{\infty} c_n \sum_{m=0}^{n} \binom{n}{m} (x-a)^m a^{n-m} \\ &= \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} c_n \binom{n}{m} a^{n-m} (x-a)^m \end{aligned}$$

put $d_m = \sum_{n=m}^{\infty} c_n \binom{n}{m} a^{n-m}$ are done. So all we need to justify is that the third equality. Indeed from the previous proposition if we have

$$\sum_{n=0}^{\infty} \sum_{m=0}^{n} |c_n| \binom{n}{m} |x-a|^m |a|^{n-m} < \infty$$

which is the same as

$$\sum_{n=0}^{\infty} |c_n| (|x-a| + |a|)^n$$

the same c_n coefficients, so radius doesn't change, we need $|x-a|+|a| < R$ or $|x-a| < R - |a|$. ☐

Theorem 51. *Ω open, connected in $\mathbb{C}$ or $\mathbb{R}$ and f is analytic on Ω then $Z(f)$, set of zeros, has no limit point unless $f = 0$.*

Proof. (the proof is the same for $\mathbb{C}$ or $\mathbb{R}$, we follow from Rudin II)

Note that $Z(f)$ is closed ($\because Z(f) = f^{-1}(0)$ and f is C^∞), let $A =$ limit points of $Z(f)$, so $A \subset Z(f)$.

$\forall\, a \in A$, $\exists\, B_{a,R} \subset \Omega$ s.t.

$$f(z) = \sum_{n=0}^{\infty} c_n (z-a)^n \qquad \forall\, z \in B_{a,R}$$

and $f(a) = 0 \implies c_0 = 0 \implies$ either $\exists$ smallest m s.t. $c_m \neq 0$, or $c_m = 0 \;\forall\, m \geq 0$.

In the second case $f(z) = 0 \;\forall\, z \in B_{a,R} \implies B_{a,R} \subset A \implies a$ is an interior point of $A \implies A$ is open.

In the first case, $f(z) = \sum_{n=m}^{\infty} c_m (z-a)^n$, $c_m \neq 0$, so

$$f(z) = (z-a)^m \sum_{n=0}^{\infty} c_{n+m}(z-a)^n$$

set $g(z) = \sum_{n=0}^{\infty} c_{n+m}(z-a)^n$, then g is analytic in $B_{a,R}$ and $g(a) \neq 0$ ($\because g(a) = c_m$), so $\exists\, \rho > 0$ s.t.

$$g(z) \neq 0 \quad \forall\, z \in B_{a,\rho}$$

$f(z) \neq 0\ \forall\, z \in B_{a,\rho} - \{a\} \implies a$ is a isolated zero point, i.e. $a \notin A \implies A = \varnothing$. So we are back to the second case.

Let $B = \Omega - A$, we will show B is open. Suppose the other way, if $b \in B$ and b is not an interior point in B, then $\forall\, r$

$$B_{b,r} \cap A \neq \varnothing$$

So b is a limit point of A, so $b \in A$ contradiction, so B is open, but Ω is connected, so $B = \varnothing \implies f = 0$, or $A = \varnothing$. $\square$

En route we have proved

Proposition 52. *If a is an isolated zero, $\exists!\, m \in \mathbb{N}$ s.t.*

$$f(z) = (z-a)^m g(z)$$

where $g(z)$ is analytic and $g(a) \neq 0$, m is called order of zero of a.

The converse of the proposition is not true. With little modification, one can use this to prove for Laurent series in complex analysis.

Lecture 7 (2/12/13)

Corollary 53. *f_1, f_2 analytic on Ω open, connected and $f_1(z) = f_2(z)$ $\forall\, z \in E \subset \Omega$, E has a limit point, then*

$$f_1 = f_2$$

Proof. use $f_1 - f_2$. $\square$

This gives uniqueness of analytic continuation. If f is analytic on Ω_1 and g, extension of f, is analytic on $\Omega_2 \supset \Omega_1$, i.e.

$$g|_{\Omega_1} = f$$

and g is unique. This is very surprising result, because for normal C^∞ functions, there are ∞ many extensions.

Definition 54. X metric space is σ compact iff X is the union of countably many compact subspaces.

$\mathbb{R}^n$, $\mathbb{C}^n$ are σ compact.

Corollary 55. *An non zero analytic function on* $\mathbb{R}^n$, $\mathbb{C}^n$ *has at most countably many zeros.*

Proof. Since the space is σ compact, we can cover the domain Ω with countably many compact sets. An infinite set in a compact has a limit point, so $Z(f)$ has finitely many points in each compact and a countable union of finite is countable. $\square$

2.2 Fourier Series

We do this from functional perspective. Let's start with some big pictures.

Definition 56. Define l^2 space

$$l^2 = \left\{ \text{complex sequences } \{c_n\}_{n \in \mathbb{N}} \text{ s.t. } \sum_{n=0}^{\infty} |c_n|^2 < \infty \right\}$$

Define a distance on l^2,

$$\|\{c_n\}_{n\in\mathbb{N}}\|_{l^2} = \sqrt{\sum_{n=0}^{\infty} |c_n|^2}$$

check this defines a metric space. This is in fact not only a metric space, but also a ∞ dimensional vector space over $\mathbb{C}$. One possible basis

$$e_i = (0, 0, ..., 0, 1, 0, ...)$$

We have dot product

$$\{a_n\}_{n\in\mathbb{N}} \cdot \{b_n\}_{n\in\mathbb{N}} = \sum_{n=0}^{\infty} a_n \bar{b}_n$$

which is compatible with norm

$$\{a_n\} \cdot \{a_n\} = \|a_n\|_{l^2}^2$$

Recall:

Definition 57. dot product $\langle f, g\rangle$ satisfies

1) linear in f, $\langle af_1 + bf_2, g\rangle = a\langle f_1, g\rangle + b\langle f_2, g\rangle$

2) anti linear in g, $\langle f, ag_1 + bg_2\rangle = \bar{a}\langle f, g_1\rangle + \bar{b}\langle f, g_2\rangle$

3) $\langle f, g\rangle = \overline{\langle g, f\rangle}$

4) $\langle f, f\rangle = \|f\|^2$ (one can defined dot product with norm, but once dot product is defined the norm comes for free.)

5) $|\langle f, g\rangle| \le \|f\|\,\|g\|$

Define L^2 space

Definition 58. let $f : [a, b] \to \mathbb{C}$

$$L^2([a, b]) = \left\{ f(x) \text{ on } [a, b] \text{ complex valued s.t. } f(x) \in \mathcal{R} \text{ and } \int_a^b |f(x)|^2 dx < \infty \right\}$$

Normally $f(x) \in \mathcal{R} \implies \int_a^b |f(x)|^2 dx < \infty$, the reason we still emphasis $\int_a^b |f(x)|^2 dx < \infty$ is to exclude improper Riemann integral

Example 59. $f(x)$, $g(x)$ on $[0, 1]$

$$f(x) = \begin{cases} \frac{1}{\sqrt{x}} & x \neq 0 \\ 0 & x = 0 \end{cases} \notin L^2 \quad g(x) = \begin{cases} \frac{1}{\sqrt[4]{x}} & x \neq 0 \\ 0 & x = 0 \end{cases} \in L^2$$

Define norm

$$\|f\|_{L^2} = \sqrt{\int_a^b |f(x)|^2 dx}$$

Note 60. Strictly speaking this is not a norm, because $\|f\|_{L^2} = 0 \not\Longrightarrow f = 0$, therefore we add an equivalence $\sim$

$$f \sim g \iff \|f - g\|_{L^2} = 0$$

and $L^2/\sim$ is a metric space we actually study, but we still write it as L^2.

Question 61. *Why not define L^2 to be continuous function?*

This will make too restricted, because we want to make connection between L^2 and l^2.

We define dot product on L^2,

$$\langle f, g\rangle = \int_a^b f(x)\bar{g}(x)dx$$

Exercise 62. Check this is indeed a dot product. Hint: use the following inequality.

Proposition 63. *(Hölder, ö sounds /eh/)* $p, q > 0$ $p, q \in \mathbb{R}$ *s.t.* $\frac{1}{p} + \frac{1}{q} = 1$, *then*

$$\left| \int_a^b f \bar{g} dx \right| \leq \left(\int_a^b |f|^p dx \right)^{\frac{1}{p}} \left(\int_a^b |g|^q dx \right)^{\frac{1}{q}}$$

That $p = q = 2$ gives Cauchy-Schwartz.

Question: we have a basis for l^2, can we find a basis for L^2? then we might have a isomorphism (or even isometry) from basis of l^2 to basis of L^2:

$$\begin{aligned} \phi : L^2([a, b]) &\rightarrow l^2 \\ f(x) &\mapsto \{c_n\}_{n \in \mathbb{N}} \end{aligned} \tag{2.1}$$

Definition 64. $\varphi_n(x) \in L^2$ are an orthonormal system iff

$$\langle \varphi_n, \varphi_m \rangle = \delta_{nm} = \begin{cases} 1 & n = m \\ 0 & n \neq m \end{cases}$$

δ_{nm} called Kronecker delta.

Recall if $\varphi_n(x)$ is a basis of L^2, this means $\forall\, f \in L^2$

$$f(x) = \sum_{n=0}^{\infty} c_n \varphi_n(x) \text{ and } c_n \in \mathbb{C} \text{ are unique}$$

If everything we wrote down makes sense, then

$$c_n = \langle f, \varphi_n \rangle$$

and ϕ of equation (2.1) is

$$\phi : f(x) \mapsto \phi(f) = c_n = \langle f, \varphi_n \rangle$$

We would like to have

(1) ϕ is an isomorphism of vector spaces

(2) ϕ is an isometry, i.e.

$$\|f\|_{L^2} = \|\phi(f)\|_{l^2}$$

That is

$$\int_a^b |f|^2 dx = \sum_{n=0}^{\infty} |c_n|^2$$

Remark 65. (1) is not true; ϕ is not surjective. However (1) will be true if we replace Riemann integral by Lebesgue integral. The reason is that l^2 is complete, but L^2 is not complete in Riemann but becomes complete in Lebesgue. E.g. Dirichlet function is Lebesgue not Riemann.

(2) is true.

Let's now study Fourier formally.

Proposition 66. *$f \in L^2([a,b])$ and φ_n is an orthonormal system, (we don't know whether it defines a basis), put*

$$S_N(f) = \sum_{n=0}^{N} c_n \varphi_n(x)$$

where $c_n = \langle f, \varphi_n \rangle$. Let $d_n \in \mathbb{C}$ be any number, let

$$T_N = \sum_{n=0}^{N} d_n \varphi(x)$$

then

$$\|f - S_N(f)\|_{L^2} \leq \|f - T_N\|_{L^2}$$

and we have equality iff $c_n = d_n \; \forall \, n = 0, 1, ..., N$.

This says $S_N(f)$ is a better approximation then T_N in L^2.

Proof.

$$\begin{aligned} \|f - T_N\|^2 &= \langle f - T_N, f - T_N \rangle \\ &= \|f\|^2 + \|T_N\|^2 - \langle T_N, f \rangle - \langle f, T_N \rangle \end{aligned}$$

$$\begin{aligned} \|T_N\|^2 &= \int_a^b \sum_{n=0}^{N} \sum_{m=0}^{N} d_n \bar{d}_m \varphi_n(x) \bar{\varphi}_m(x) dx \\ &= \sum_{n,m} d_n \bar{d}_m \langle \varphi_n, \varphi_m \rangle \;\because \text{finite sum can swap} \\ &= \sum_{n=0}^{N} |d_n|^2 \end{aligned}$$

$$\langle f, T_N \rangle = \int_a^b f(x) \sum_{n=0}^{N} \bar{d}_n \bar{\varphi}_n dx = \sum_{n=0}^{N} \bar{d}_n c_n$$

hence

$$\|f - T_N\|_{L^2}^2 = \|f\|^2 + \sum |d_n|^2 - \sum \bar{c}_n d_n - \sum c_n \bar{d}_n = \|f\|_{L^2}^2 + \|T_N - S_N\|_{l^2}^2 - \sum |c_n|^2$$

Set $T_N = S_N$ above, we get

$$\|f - S_N\|_{L^2}^2 = \|f\|_{L^2}^2 - \sum |c_n|^2$$

This proves the statement. □

Corollary 67. *(Bessel)Let $c_n = \langle f, \varphi_n \rangle$, then*

$$\sum_{n=0}^{N} |c_n|^2 \leq \|f\|_{L^2}^2$$

and if $f \in L^2$, then

$$\sum_{n=0}^{\infty} |c_n|^2 < 0 \text{ and } c_n \to 0 \text{ as } n \to \infty$$

Example 68. One possible form of φ_n is for $f \in L^2([-\pi, \pi])$ is periodic on $[-\pi, \pi]$ (call f is defined on a circle), put

$$\varphi_n(x) = \frac{e^{inx}}{\sqrt{2\pi}}$$

$$S_N(f) = \sum_{n=-N}^{N} c_n \frac{e^{inx}}{\sqrt{2\pi}} \qquad c_n = \int_{-\pi}^{\pi} f(x) \frac{e^{-inx}}{\sqrt{2\pi}} dx \tag{2.2}$$

Lecture 8 (2/14/13)

Remark 69. (continuous) periodic functions are in bijection with (continuous) periodic functions on a circle.

Proof. Let

$$S = \mathbb{R}/\sim \quad x \sim y \iff x - y = 2\pi n \;\; n \in \mathbb{Z}$$

S^1 with quotient topology is coarsest topology on circle, and the map is continuous

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{q} & S^1 \\ x & \mapsto & [x] = \{x + 2\pi n\} \end{array}$$

by checking open per-image are open in $\mathbb{R}$.

Consider $f : \mathbb{R} \to \mathbb{R}$ or $\mathbb{C}$, $f = \tilde{f} \circ q$

$$\mathbb{R} \xrightarrow{q} S^1 \xrightarrow{\tilde{f}} \mathbb{R} \text{ or } \mathbb{C}$$

so that

$$\tilde{f}([x]) = f(x)$$

For $\tilde{f}$ to exist, we need

$$x \sim y \text{ s.t. } q(x) = q(y) \implies f(x) = f(y)$$

that is f is periodic. This is satisfied.

Now show $\tilde{f}$ is continuous $\iff$ f is continuous.

Clearly if $\tilde{f}$ is continuous, then $f = \tilde{f} \circ q$ is continuous.

Conversely V open in $\mathbb{R}$ or $\mathbb{C}$,

$$q^{-1}(\tilde{f}^{-1}(V)) = f^{-1}(V)$$

which is open, for f is continuous. Since S^1 has quotient topology, $\tilde{f}^{-1}(V)$ is open, so $\tilde{f}$ is continuous. $\square$

Advantage of $\tilde{f}$ over f is that $\tilde{f}$ is on S^1 which is compact.

Example 70. (Dirichlet Kernel) From (2.2),

$$\begin{aligned}
S_N(f)(x) &= \sum_{-N}^{N} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt e^{inx} \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \sum_{-N}^{N} e^{in(x-t)} dt \\
&= \frac{1}{2\pi} \int_{x-\pi}^{x+\pi} f(x-u) \sum_{-N}^{N} e^{inu} du \text{ set } t = x - u \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-u) \left(\sum_{-N}^{N} e^{inu} \right) du \because \text{both have period } 2\pi \\
&= f * D_N(x)
\end{aligned}$$

where

$$D_N(u) = \frac{1}{2\pi} \sum_{-N}^{N} e^{inu}$$

Find $D_N(u)$

$$e^{iu}D_N - D_N = \frac{e^{i(N+1)u} - e^{-iNu}}{2\pi}$$

So

$$D_N = \frac{1}{2\pi}\frac{e^{i(N+\frac{1}{2})u} - e^{-i(N+\frac{1}{2})u}}{e^{i\frac{u}{2}} - e^{-i\frac{u}{2}}} = \frac{1}{2\pi}\frac{\sin(N+\frac{1}{2})u}{\sin\frac{u}{2}}$$

Definition 71. Functions f on D is Lipschitz iff

$$|f(x) - f(y)| \leq M|x-y|$$

$\forall\, x, y \in D$.

One can replace $|\quad|$ by $\|\quad\|$ to make it work on metric spaces.

Remark 72. Lipschitz $\implies$ uniform continuous $\implies$ continuous.

Proposition 73. $f(t) \in L^2([-\pi,\pi])$ *and Lipschitz in a neighborhood of* x *i.e.*

$$|f(t) - f(u) \leq M|t-u| \quad \text{if } t, u \in B_{x,\delta}$$

then

$$S_N(t)(x) \to f(x) \text{ as } N \to \infty$$

Most of time one gets pointwise convergence (as stated above) but not uniform convergence, however always converges in L^2, cf Parseval theorem.

Proof. We use the similar trick in proof of Abel

Because $\int_{-\pi}^{\pi} D_N(t)dt = \sum_{-N}^{N}\int_{-\pi}^{\pi}\frac{e^{int}}{2\pi}dt = 1$,

$$\begin{aligned}
f(x) - S_N(f)(x) &= f(x) - \int_{-\pi}^{\pi} f(x-t)D_N(t)dt \\
&= \int_{-\pi}^{\pi}(f(x) - f(x-t))\,D_N(t)dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}(f(x) - f(x-t))\left(\sin Nt\frac{\cos\frac{t}{2}}{\sin\frac{t}{2}} + \cos Nt\right)dt
\end{aligned}$$

Put $g(x) = (f(x) - f(x-t))\,\frac{\cos\frac{t}{2}}{\sin\frac{t}{2}}$. If $g(x) \in L^2$, by corollary 67, $\sin Nt$ are orthogonal systems

$$\int_{-\pi}^{\pi}(f(x) - f(x-t))\,\frac{\cos\frac{t}{2}}{\sin\frac{t}{2}}\sin Nt dt = \langle g(x), \sin Nt\rangle \to 0$$

Indeed if $|t| < \delta$, because of Lipschitz, $|f(x) - f(x-t)| \leq M|t|$, then

$$\left|(f(x) - f(x-t))\frac{\cos\frac{t}{2}}{\sin\frac{t}{2}}\right| \leq M\left|\frac{t}{\sin\frac{t}{2}}\right| |\cos\frac{t}{2}|$$

both terms on the right go to 1 as $t \to 0$, hence if $|t| < \delta$, RHS is small.

If $|t| > \delta$, $\sin\frac{t}{2}$ is not too small, so $g(x)$ is too bounded. Therefore $g \in L^2$.

Similarly

$$\int_{-\pi}^{\pi} (f(x) - f(x-t)) \cos Nt dt \to 0$$

$\square$

Remark 74. $f \in L^2$ s.t. $f(x) = 0\ \forall\, x \in (a,b) \subset [-\pi,\pi]$, by the proposition

$$S_N(f)(x) \to 0$$

$\forall\, x \in (a,b)$.

Proposition 75. *$f(x) \in C([-\pi,\pi])$ s.t. $f(\pi) = f(-\pi)$, let $A = \{\sum_{-N}^{N} c_n e^{inx}\}$ sub algebra of $C([-\pi,\pi])$, then*

$$\bar{A} = C([-\pi,\pi])$$

$\bar{A}$ is the uniform closure, or call $\|\ \|_\infty$ closure. it means

$$\forall\, \epsilon, \exists\, c_n \text{ s.t. } \forall\, x \quad \left|f(x) - \sum_{-N}^{N} c_n e^{inx}\right| < \epsilon$$

Proof. A separates points (use $\sin x, \cos x \in A$) and doesn't vanish ($x = 0 \implies$ constant function $\in A$) and self-adjoint

Use Stone on $C(S^1)$, get $\exists\, f \in S^1$

$$\left|f(x) - \sum_{-N}^{N} c_n e^{inx}\right| < \epsilon$$

Then use bijection between continuous on S^1 and periodic continuous function on $\mathbb{R}$, cf remark 69. $\square$

Theorem 76. *(Parseval) $f \in L^2([-\pi,\pi])$ periodic*

1)

$$\|S_N(f) - f\|_{L^2} \to 0$$

2) $S_N(f) = \sum c_n e^{inx}$, $S_N(g) = \sum d_n e^{inx}$,

$$\int_{-\pi}^{\pi} f(x)\bar{g}(x)dx = \sum_{-\infty}^{\infty} c_n \bar{d}_n$$

3)

$$\begin{array}{rcl} L^2 & \to & l^2 \\ f & \mapsto & \{c_n\}_{n\in\mathbb{N}} \end{array}$$

is an isometry, i.e.

$$\|f\|_{L^2}^2 = \sqrt{\sum |c_n|^2}$$

This shows that two norms are the same.

Exercise 77. (chapter 6 exercise 12) Continuous functions are dense in L^2.

That is $\forall\, \epsilon > 0\ \exists\, g \in C([-\pi,\pi])$ s.t. $\|f - g\|_{L^2} < \epsilon$, or the $\|\ \|_{L^2}$ closure $\overline{C([-\pi,\pi])} = L^2([-\pi,\pi])$.

Proof. 1) Fix $\epsilon > 0$ use exercise to find $g \in C([-\pi,\pi])$, then use Stone on g to find $P(x) = \sum_{-N_0}^{N_0} d_n e^{inx}$,

$$|g(x) - P(x)| < \epsilon \quad \forall\, x \in [-\pi,\pi]$$

After integrate, we obtain

$$\|g(x) - P(x)\|_{L^2} \leq \sqrt{2\pi}\epsilon$$

By proposition 66

$$\|g - S_N(g)\|_{L^2} \leq \|g - P\|$$

Therefore

$$\begin{aligned} \|f - S_N(f)\|_{L^2} &\leq \|f - g\| + \|g - S_N(g)\| + \|S_N(g) - S_N(f)\| \\ &\leq \epsilon + \sqrt{2\pi}\epsilon + \epsilon \end{aligned}$$

$\|S_N(g) - S_N(f)\| = \|S_N(g - f)\| \leq \|g - f\|$ by corollary 67.

Lecture 9 (2/19/13)

2)

$$\begin{aligned}
\int_{-\pi}^{\pi} S_N(f)\bar{g}dx &= \sum_{-N}^{N}\int_{-\pi}^{\pi} c_n \frac{e^{inx}}{\sqrt{2\pi}}\bar{g}dx \\
&= \sum_{-N}^{N} c_n \overline{\int_{-\pi}^{\pi} g\frac{e^{-inx}}{\sqrt{2\pi}}dx} = \sum_{-N}^{N} c_n\bar{d}_n
\end{aligned}$$

Since

$$\begin{aligned}
\left|\sum_{-N}^{N} c_n\bar{d}_n - \int f\bar{g}dx\right| &= \left|\int_{-\pi}^{\pi}\left(S_N(f)\bar{g} - f\bar{g}\right)dx\right| \\
&\leq \int_{-\pi}^{\pi} |S_N(f) - f|\,|g|dx \\
&\leq \|S_N(f) - f\|_{L^2}\,\|g\|_{L^2} \text{ Cauchy-Schwartz}
\end{aligned}$$

By 1) $\|S_N(f) - f\|_{L^2} \to 0$. Hence if $N \geq N_\epsilon$,

$$\left|\sum_{-N}^{N} c_n\bar{d}_n - \int f\bar{g}dx\right| < \epsilon$$

this is not an equivalent norm, it is the ordinary norm in $\mathbb{R}$, so

$$\sum_{-N}^{N} c_n\bar{d}_n = \int f\bar{g}dx$$

3) set $f = g$ in 2) $\square$

So we don't have distinguish $\| \|_{L^2}$ and $\| \|_{l^2}$, and we'll just call it $\| \|_2$. We'll go over Parseval again after we have Lebesgue.

2.3 Gamma Functions

A few words about improper integrals,

$$\int_0^\infty f(x)dx = \lim_{M\to\infty}\int_0^M f(x)dx$$

$$\int_{-\infty}^\infty f(x)dx = \lim_{M\to\infty}\int_{-M}^M f(x)dx$$

Lebesgue will already define on $\int_{-\infty}^{\infty} f(x)dx$, but there are cases that Riemann exists not Lebesgue, e.g.

$$\int_{-\infty}^{\infty} \frac{\sin x}{x} dx$$

Definition 78. Γ function

$$\Gamma(x) = \int_{0}^{\infty} t^{x-1} e^{-t} dt$$

domain $(0, \infty)$.

If x is very small positive or negative (no necessary small), both t^{x-1} and e^{-t} go to 0 quickly, so the integral is roughly

$$\Gamma(x) \approx \lim_{\epsilon \to 0^+} \int_{\epsilon}^{M} t^{x-1} e^{-t} dt$$

where M is not too big. In this interval, $e^{-t} \approx 1$, so we do

$$\Gamma(x) \approx \lim_{\epsilon \to 0^+} \int_{\epsilon}^{M} t^{x-1} dt = \lim_{\epsilon \to 0^+} \left. \frac{t^x}{x} \right|_{\epsilon}^{M} < \infty \iff x > 0$$

This shows the natural domain of Γ is $x > 0$.

To show integral converges for x near ∞ use the following induction.

Proposition 79. $\Gamma(x+1) = x\Gamma(x)$.

Proof.

$$\Gamma(x+1) = -t^x \, e^{-t} \big|_0^{\infty} + x \int_0^{\infty} t^{x-1} e^{-t} dt = x\Gamma(x).$$

□

Since $\Gamma(1) = 1$, $\Gamma(n+1) = n! \;\; n \in \mathbb{N}$.

Proposition 80. *$\Gamma(x)$ is convex and $\log \Gamma(x)$ is also convex.*

Proof. That $f(x)$ is convex means for x, y, on the segment $(1-t)x + ty \qquad 0 \le t \le 1$,

$$f\left((1-t)x + ty\right) \le (1-t)f(x) + tf(y)$$

set $t = \frac{1}{q}$, so $q > 1$, $1 - t = \frac{1}{p}$

$$\begin{aligned}
\Gamma\left((1-t)x + ty\right) &= \Gamma\left(\frac{x}{p} + \frac{y}{q}\right) \\
&= \int_0^\infty t^{\frac{x}{p}+\frac{y}{q}-1} e^{-t} dt \\
&= \int_0^\infty \left(t^{\frac{x-1}{p}} e^{-\frac{t}{p}}\right)\left(t^{\frac{y-1}{q}} e^{-\frac{t}{q}}\right) dt \because \text{Holder} \\
&\leq \Gamma(x)^{\frac{1}{p}} \Gamma(x)^{\frac{1}{q}} \qquad (2.3) \\
&\leq \frac{1}{p}\Gamma(x) + \frac{1}{q}\Gamma(x) \because \text{Young inequality}
\end{aligned}$$

showing $\Gamma(x)$ is convex.

Take log of (2.3), since log is increasing

$$\log \Gamma\left(\frac{1}{p}x + \frac{1}{q}y\right) \leq \frac{1}{p}\Gamma(x) + \frac{1}{q}\Gamma(y)$$

□

Note 81. Young inequality: $a^{\frac{1}{p}} b^{\frac{1}{q}} \leq \frac{1}{p}a + \frac{1}{q}b$, or if put $\tilde{a} = a^{\frac{1}{p}}$, $\tilde{b} = b^{\frac{1}{q}}$, $\Longrightarrow$ alternative form of Young

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

Remark 82. So people call $\Gamma(x)$ super convex. In general

$$\log f(x) \text{ is convex} \implies f(x) \text{ is convex}$$

Proof: $\log f(x)$ is convex $\implies e^{\log f(x)} = f(x)$ is convex, for exponential is convex and composition of two convex functions is convex.

We want study the asymptotic behavior of Γ near ∞. Hence we want to find $g(x)$, s.t.

$$\lim_{x\to\infty} g(x)\Gamma(x) = 1$$

Since

$$\begin{aligned}
\Gamma(x+1) &= \int_0^\infty e^{x\log t - t} dt \\
&= x\int_0^\infty e^{x\log u + x\log x - ux} du \text{ set } t = ux \\
&= x^{x+1}\int_0^\infty e^{x(\log u - u)} du \\
&= x^x \int_0^\infty e^{x(\log u - u)} du \qquad (2.4)
\end{aligned}$$

We notice function $\log u - u$ goes up and down for $u \in (0, \infty)$, and have maximum at $u = 1$. This leads us to study a more general situation:

Theorem 83. *(Laplace's method) We want to solve the following problem: Given $h(t)$, s.t.*

1) $h(t) \in C^2(0, \infty)$

2) $h(t) \to -\infty$ as $t \to 0^+$ and $t \to \infty$

3) $\exists$ only one t_0 s.t. $h'(t_0) = 0$

4) $h''(t_0) < 0$

we would like to find $g(x)$ s.t.

$$\lim_{x \to \infty} g(x) \int_0^\infty e^{xh(t)} dt = 1$$

then

$$g(x) = e^{-ax} \sqrt{\frac{bx}{2\pi}}$$

Proof. (Rudin proof is a bit cheat. It uses Lebesgue to move limit inside of integral. The proof we follow is from book by Dieudonné. The book is full of big notations, very French style)

Consider x big. We break $t > 0$ into $(0, t_0 - \epsilon)$, $(t_0 - \epsilon, t_0 + \epsilon)$, and $(t_0 + \epsilon, \infty)$

If $t \in (t_0 - \epsilon, t_0 + \epsilon)$, we do Taylor for $h(t)$, not for $e^{xh(t)}$, because $e^{xh(t)}$ grows too fast, Taylor is not very accurate.

$$h(t) = h(t_0) + \frac{h''(t_0)}{2}(t - t_0)^2 + O\left((t - t_0)^3\right)$$

where O is Landau big O notation, meaning $f(t)$ is $O(g(t))$ at t_0 iff $\exists\, \epsilon, \delta, M$ s.t.

$$t_0 - \epsilon < t < t_0 - \delta \implies |f(t)| \leq M|g(t)|$$

Let $h(t_0) = a$, $h''(t_0) = -b$ $(b > 0)$

$$e^{xh(t)} = e^{xa} e^{-\frac{xb}{2}(t - t_0)^2} \left(1 + xO\left((t - t_0)^3\right)\right)$$

because $e^{xO\left((t-t_0)^3\right)} \approx 1 + xO\left((t-t_0)^3\right)$ for $|t-t_0|$ small, so

$$\begin{aligned}
\int_{-\epsilon+t_0}^{\delta+t_0} e^{xh(t)}dt &= e^{xa}\int_{-\epsilon}^{\delta} e^{-\frac{xb}{2}t^2}dt + xe^{xa}\int_{-\epsilon}^{\delta} e^{-\frac{xb}{2}t^2}O(t^3)dt \\
&= e^{xa}\sqrt{\frac{2}{bx}}\underbrace{\int_{-\epsilon\sqrt{\frac{bx}{2}}}^{\delta\sqrt{\frac{bx}{2}}} e^{-u^2}du}_{\substack{\to\sqrt{\pi} \\ \text{as } x\to\infty}} + ...
\end{aligned}$$

Hence

$$\lim_{x\to\infty}\sqrt{\frac{bx}{2\pi}}\int_{-\epsilon}^{\delta} e^{-\frac{xb}{2}t^2}dt = 1$$

What about the O part?

$$\left|xe^{xa}\int_{-\epsilon}^{\delta} e^{-\frac{xb}{2}t^2}O(t^3)dt\right| \leq xe^{xa}M\int_{-\epsilon}^{\delta} e^{-\frac{xb}{2}t^2}t^3dt \sim e^{-\frac{xb}{2}\delta^2} \to 0 \text{ as } x\to\infty$$

Hence

$$\lim_{x\to\infty} e^{-ax}\sqrt{\frac{bx}{2\pi}}\int_{-\epsilon}^{\delta} e^{xh(t)}dt = 1$$

Now do $t \in (0, t_0-\epsilon)$, cannot use Taylor. $h(t) \to -\infty$, Fix $\epsilon > 0$ s.t. $h(t) < -1$ for $0 < t < t_0 - \epsilon$, Assume $x > 1$,

$$(x-1)h(t) < -(x-1) \implies e^{xh(t)} < e^{-(x-1)+h(t)}$$

$$0 \leq \int_0^{t_0-\epsilon} e^{xh(t)}dt \leq e^{-(x-1)}\int_0^{t_0-\epsilon} e^{h(t)} \to 0$$

Hence

$$\lim_{x\to\infty}\int_0^{t_0-\epsilon} e^{xh(t)}dt = 0$$

Similarly show $\lim_{x\to\infty}\int_{t_0+\epsilon}^{\infty} e^{xh(t)}dt = 0$, therefore

$$\lim_{x\to\infty} e^{-ax}\sqrt{\frac{bx}{2\pi}}\int_0^{\infty} e^{xh(t)}dt = 1$$

$\square$

Apply Laplace's method to (2.4), $t_0 = 1$, $a = \log 1 - 1 = -1$, $b = (\log u - u)''_{u=1} = -1$, then as $x \to \infty$

$$\frac{e^x \Gamma(x+1)}{x^x \sqrt{2\pi x}} \longrightarrow 1$$

If $x \in \mathbb{N}$,

$$\frac{e^n n!}{n^n \sqrt{2\pi n}} \longrightarrow 1$$

Hence we derived

Proposition 84. *(Stirling)*

$$n! \approx n^n \sqrt{2\pi n} e^{-n}$$

as $n \to \infty$.

The following proposition is ridiculously challenging to prove (proved by analytic continuation, uniform limit of analytic functions), so we will just state the result.

Proposition 85. Γ *function is analytic everywhere.*

Idea of proof: One can define Γ for negative $\mathbb{R}$, except $\{0, -1, -2, ...\}$. First show Γ is analytic for $(0, +\infty)$. Use analytic continuation,

$$\frac{\Gamma(x+1)}{x} =: \Gamma(x) \quad \frac{\Gamma(x+2)}{(x+1)x} =: \Gamma(x)$$

for $\mathbb{C}$, and check $x = 0, -1, -2, ...$ are poles of degree 1.

Last result of Γ

Lecture 10 (2/26/13)

Theorem 86. $\phi(x)$ *s.t.*

$$\left\{ \begin{array}{c} \phi(x+1) = x\phi(x) \\ \phi(1) = 1 \end{array} \right\} \implies \{\phi(n+1) = n!\}$$

and $\log \phi(x)$ *is convex, then*

$$\phi(x) = \Gamma(x)$$

We need the following lemma, which is also the preliminary lemma for proving second derivative of convex function exists, and it is positive.

Lemma 87. *$f(x)$ is convex, then, let $t < u < v$,*

$$\frac{f(u)-f(t)}{u-t} \leq \frac{f(v)-f(t)}{v-t} \leq \frac{f(v)-f(u)}{v-u}$$

Proof. (of theorem) Suppose $0 < x < 1$, apply lemma for $n < n+1 < n+1+x < n+2$

$$\begin{aligned} \log n = \frac{\log\phi(n+1)-\log\phi(n)}{1} &\leq \frac{\log\phi(n+1+x)-\log\phi(n+1)}{x} \\ &\leq \frac{\log\phi(n+2)-\log\phi(n+1)}{1} = \log(n+1) \end{aligned}$$

Since $\phi(n+1+x) = (n+x)\cdot ... \cdot x\phi(x)$,

$$\log n^x \leq \log\phi(x) + \log\frac{(n+x)\cdot ... \cdot x}{n!} \leq \log(n+1)^x$$

then subtract $\log n^x$

$$0 \leq \log\phi(x) + \log\frac{(n+x)\cdot ... \cdot x}{n!n^x} \leq \log(1+\frac{1}{n})^x$$

this works $\forall\, n$. Take $n \to \infty$,

$$\log\phi(x) = \lim_{n\to\infty} \log\frac{n!n^x}{(n+x)\cdot ... \cdot x}$$

That log is continuous implies

$$\phi(x) = \lim_{n\to\infty} \frac{n!n^x}{(n+x)\cdot ... \cdot x}$$

showing uniqueness. But we always show such ϕ exists which is Γ, so

$$\Gamma(x) = \lim_{n\to\infty} \frac{n!n^x}{(n+x)\cdot(n-1+x)... \cdot x}$$

check this formula works $\forall\, x > 0$. $\square$

3 Muti-variables Calculus

Consider $\Omega \subset \mathbb{R}^m$ open,

$$f : \Omega \to \mathbb{R}^n$$

We want to a concept of derivatives

$$f'(x_0) : \mathbb{R}^m \to \mathbb{R}^n$$

$x_0 \in \Omega$ and $f'(x_0)$ satisfies

$$\lim_{h\to 0} \frac{|f(x_0+h) - f(x_0) - f'(x_0)h|}{|h|} = 0 \tag{3.1}$$

It turns out that $f'(x)$ is $n \times m$ matrix, the limit is over any vector $\vec{h}$ (in any direction), while the divisor is the length. We call $f'(x)$ the first derivative of $f(x)$.

Definition 88. If $\exists\, f'(x_0) : \mathbb{R}^m \to \mathbb{R}^n$ linear s.t. the limit (3.1) satisfied, we say that f is differentiable at x_0 and $f'(x_0)$ is the derivative.

Note 89. The way we define higher dimensional is defined here, (3.1), even in 2 dimension ($n = m = 2$) is not the complex derivative, there is a theorem says that under some condition it becomes equivalent to complex derivative.

3.1 Functions of Several Variables

Let's review linear algebra.

Definition 90. K field, V a set is called a $K-$vector space iff two operations

$$V \times V \xrightarrow{+} V \qquad K \times V \xrightarrow{\cdot} V$$

s.t. $(V, +)$ is a commutative group. We need to have compatibility of two operations with the $(+, \cdot)$ operation on K

1) distributive $(a+b)\vec{v} = a\vec{v} + b\vec{v}$, $a(\vec{v}_1 + \vec{v}_2) = a\vec{v}_1 + a\vec{v}_2$

2) associative $(ab)\vec{v} = a(b\vec{v})$

3) neutral element $1 \in K$ $1 \cdot \vec{v} = \vec{v}$

One can define module over vector space indeed of field, so that it becomes less restricted.

Definition 91. $\vec{v}_1, ..., \vec{v}_n \in V$. They are linearly independent iff

$$\sum_{i=1}^{n} a_i \vec{v}_i = 0 \implies a_i = 0 \; \forall\, i$$

Definition 92. $\langle\vec{v}_1, ..., \vec{v}_n\rangle$ the span of the n vectors is the set of all linear combinations,

$$\langle\vec{v}_1, ..., \vec{v}_n\rangle = \{\sum_{i=1}^{n} a_i\vec{v}_i, a_i \in K\}$$

Definition 93. $\vec{v}_1, ..., \vec{v}_n$ are a basis of V iff they are linear independent and $\langle\vec{v}_1, ..., \vec{v}_n\rangle = V$

Proposition 94. *V a vector space with a basis $\vec{v}_1, ..., \vec{v}_n$, every basis of V has exactly n numbers, we call n the $dimV$.*

Whenever we define an algebraic structure, we also define class of functions on the structure.

Definition 95. V, W are $K-$space $f : V \to W$ is linear iff

$$\begin{cases} f(\vec{v}_1 + \vec{v}_2) = f(\vec{v}_1) + f(\vec{v}_2) \\ f(a\vec{v}) = af(\vec{v}) \end{cases}$$

When we talk about functions on vector space, they are always linear.

Corollary 96. *If V is a K-space of dimension n, V is isomorphic to K^n, n triple of K.*

The isomorphic is linear, but not unique, any map maps any basis of V to any basis of K^n. The corollary is false in ∞ dimension.

Proposition 97. *V finite dimensional vector space, $A : V \to V$ linear map*

$$A \text{ is injective} \iff A \text{ is surjective}$$

This is false in ∞ dimension. This is also false in differential geometry for a local vector space (called "vector bundle").

Proof. $\vec{v}_1, ..., \vec{v}_n$ a basis of V. Because A is linear

$$\mathrm{Im}(A) = \langle A\vec{v}_1, ..., A\vec{v}_n\rangle$$

If A is surjective, $A\vec{v}_1, ..., A\vec{v}_n$ are linearly independent, for any $\vec{v}$

$$\vec{v} = \sum a_i(A\vec{v}_i) = A\sum a_i\vec{v}_i$$

so take $\vec{v} = 0 \implies a_i = 0 \implies 0 \mapsto 0$, is injective.

If A is injective, prove that $A\vec{v}_i$ are linearly independent.

$$\sum a_i A\vec{v}_i = 0 \implies A \sum a_i \vec{v}_i = 0 \implies \sum a_i \vec{v}_i = 0 \implies a_i = 0$$

$\square$

Proposition 98. *V dim n W dim m vector spaces on K, the space of linear map*

$$L(V, W) = \{A : V \to W \ s.t \ A \ is \ linear\}$$

is a $K-$vector space of dimension nm.

$L(V, V)$ has more than $L(V, W)$, $L(V, V)$ has product: $A, B \in L(V, V)$ so is AB, which makes $L(V, V)$ be an non-commutative, associative, distributive algebra.

We now turn $L(V, W)$ into a metric space.

Definition 99. $A \in L(V, W)$ over $K = \mathbb{C}/\mathbb{R}$. define norm

$$\|A\| = \sup_{|x| \le 1} |Ax|$$

Remark 100. 1) Cauchy-Schwartz

$$|A\vec{x}| \le \|A\| \, |\vec{x}|$$

Proof: Suppose $|\vec{x}| \neq 0$, $\frac{|A\vec{x}|}{|\vec{x}|} \le \sup_{x \neq 0} \frac{|A\vec{x}|}{|\vec{x}|} = \sup_{x \neq 0} \left| A \frac{\vec{x}}{|\vec{x}|} \right| \le \sup_{|\vec{x}| \le 1} |A\vec{x}| = \|A\|$.

2)

$$|A\vec{x}| \le \lambda |\vec{x}| \ \ \forall \vec{x} \implies \|A\| \le \lambda$$

Proof: Take $\vec{x} = \vec{1}$

3)

$$\|A\| = \sup_{|\vec{x}| = 1} |A\vec{x}|$$

Proof: Let $\|A\|_1 = \sup_{x \neq 0} \frac{|A\vec{x}|}{|\vec{x}|}$

$$|Ax| \le \|A\|_1 \, |x| \ \forall x \text{ not just } |x| \le 1 \implies \|A\| \le \|A\|_1$$

On the other hand

$$\|A\|_1 = \sup_{|x|=1} |Ax| \leq \sup_{|x|\leq 1} |Ax| = \|A\|$$

4)

$$\|A\| = \max_{|x|=1} |Ax|$$

Proof: Because matrix function is continuous, norm is continuous too and on a compact space ($\because$bounded see proposition below) so sup $\implies$ max.

Lecture 11 (2/28/13)

We now prove two important propositions that are used for inverse function theorem.

Proposition 101. *1) $f(x) = Ax$ is uniformly continuous.*

2) $\|A\| < \infty$

3) $(L(V,W), \cdot)$ is metric space

4) $\|BA\| = \|B\| \, \|A\|$

Proof. 2) $\|A\| = \sup_{|x|\leq 1} |Ax|$ $e_1, ..., e_n$ a basis of V

$$|x| \leq 1 \quad x = \sum_{i=1}^{n} x_i e_i \implies |x_i| \leq 1$$

$$|Ax| = |\sum_{i=1}^{n} x_i A e_i| \leq \sum_{i=1}^{n} |x_i||Ae_i| \leq \sum_{i=1}^{n} |Ae_i| < \infty$$

That last inequality because it is a sum of n numbers that are independent of x.

1)

$$|Ax - Ay| \leq \|A\| \, |x-y| < \epsilon$$

if $|x-y| < \epsilon / \|A\|$

3) $d(A,B) = \|A - B\|$ show

(i) symmetric

(ii) $\|A\| = 0 \iff A = 0$. Indeed

$$\|A\| = 0 \implies 0 \leq |Ax| \leq 0 \cdot |x| = 0 \implies Ax = 0 \, \forall \, x$$

(iii) triangle inequality, assume $|x| \leq 1$

$$|Ax + Bx| \leq |Ax| + |Bx| \leq \|A\| + \|B\|$$

take sup over $|x| \leq 1$

$$\|A + B\| \leq \|A\| + \|B\|$$

4) Assume $|x| \leq 1$

$$|BAx| \leq \|B\| \, |Ax| \leq \|B\| \, \|A\|$$

then take sup over $|x| \leq 1$. $\square$

Proposition 102. *$L(V, V)$, $\Omega \subset L(V, V)$ is the set of invertible functions.*

1) Ω is open in $L(V, V)$

2)

$$\begin{array}{ccc} \Omega & \to & \Omega \\ A & \mapsto & A^{-1} \end{array}$$

is continuous.

This shows set of invertible matrices are open, and mapping from matrix to its inverse is continuous in this metric. So clearly Ω is not a vector space, because not closed, but it is a group with non-commutative product

$$\Omega = GL_n = GL(V)$$

general linear group.

Remark 103. 2) shows if A is continuous, then A^{-1} is continuous map. Because A^{-1} is composition of two continuous maps.

Proof. 1) Fix A invertible $A \subset \Omega$, find $r > 0$ s.t. $B_{A,r} \subset \Omega$.

We will show if $B \in L(V, V)$ s.t. $\|B - A\| < \left\|A^{-1}\right\|^{-1}$, then $B \in \Omega$, that is

$$B_{A, \|A^{-1}\|^{-1}} \subset \Omega$$

Let $\alpha = \left\|A^{-1}\right\|^{-1}$, $\beta = \|B - A\|$, so $\alpha > \beta$

$$\begin{aligned} \alpha|x| &= \alpha|A^{-1}Ax| \leq \alpha \left\|A^{-1}\right\| |Ax| \\ &\leq |(A - B)x| + |Bx| \leq \|A - B\| \, |x| + |Bx| \end{aligned}$$

So

$$\alpha|x| \leq \beta|x| + |Bx| \implies |Bx| \geq (\alpha - \beta)|x| > 0 \tag{3.2}$$

if $x \neq \vec{0}$. This shows $Bx \neq 0$, hence B is injective so B is surjective, i.e. B is invertible.

2) We show pointwise continuous, that is

$\forall\, A\ \forall \epsilon > 0\ \exists\, \delta > 0$ s.t.

$$\|A - B\| < \delta \implies \left\|A^{-1} - B^{-1}\right\| < \epsilon$$

From (3.2)

$$|BB^{-1}x| \geq (a - \beta)|B^{-1}x| \implies |B^{-1}x| \leq \frac{|x|}{\alpha - \beta}$$

then

$$\begin{aligned} \left\|A^{-1} - B^{-1}\right\| &= \left\|A^{-1}(A - B)B^{-1}\right\| \\ &\leq \left\|A^{-1}\right\| \|A - B\| \left\|B^{-1}\right\| \leq \frac{1}{\alpha}\beta\frac{1}{\alpha - \beta} = \frac{\delta}{\alpha(\alpha - \delta)} \to 0 \end{aligned}$$

□

Theorem 104. *Abstract linear map*

$$A : V \to W$$

can be written

$$Ae_i = \sum_k a_{ki} f_k$$

once we fix base. $e_1, ..., e_m$ *for* V *and* $f_1, ..., f_n$ *for* W*.* a_{ki} *is* $n \times m$ *matrix.*

Remark 105. Fix a orthonormal basis $e_1, ..., e_n$, so $|x_j| \leq |x|$,

$$\begin{aligned} |Ax| &= \left|\left(\sum_{j=1}^{n} a_{ij}x_j\right)\right| = \sqrt{\sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j\right)^2} \\ &\leq \sqrt{\sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}\right)^2}\,|x| \leq \sqrt{\sum_{i,j=1}^{m,n} |a_{ij}|^2 |x|} \end{aligned}$$

The inequalities used above can be proved using Lagrange multiplier or linear programming.

Hence

$$\|A\| \leq \sqrt{\sum_{i,j=1}^{m,n} |a_{ij}|^2}$$

Remark 106. matrix function

$$\begin{aligned} [a,b] &\to L(V,W) \\ t &\mapsto A_t = (a_{ij}(t)) \end{aligned}$$

each entry is a function

$$\begin{aligned} X &\to \mathbb{R}/\mathbb{C} \\ t &\mapsto a_{ij}(t) \end{aligned}$$

If every entry $a_{ij}(t)$ is continuous, then the matrix function is also continuous. We will use this for derivatives.

Namely $\forall\,\epsilon, \exists\,\delta > 0$ s.t.

$$d(t,s) < \delta \implies \|A_t - A_s\| < \epsilon$$

Indeed

$$\|A_t - A_s\| \leq \sqrt{\sum_{ij} |a_{ij}(t) - a_{ij}(s)|^2} < \sqrt{nm}\epsilon$$

if $|a_{ij}(t) - a_{ij}(s)| < \epsilon \Leftarrow |t-s| < \delta_{ij}$ let $\delta = \min_{i,j}\{\delta_{ij}\}$.

3.2 Derivative

Definition 107. $f : \Omega \to \mathbb{R}^n$, $\Omega \subset \mathbb{R}^m$ open, $f'(x) \in L(\mathbb{R}^m, \mathbb{R}^n)$ and the following limit exists and satisfied

$$\lim_{h\to 0} \frac{|f(x+h) - f(x) - f'(x)h|}{|h|} = 0$$

f is differentiable and f' is the derivative.

Remark 108. 1) If $\Omega \subset \mathbb{R}$, it is the old definition.

2) $\exists$ a function $r(h)$ s.t.

$$f(x+h) = f(x) + f'(x)h + r(h)$$

and

$$\lim_{h\to 0} \frac{|r(h)|}{|h|} = 0$$

where $f(x)+f'(x)h$ is called linear approximation, $r(h)$ remainder, and it goes to 0 faster than any h. We can write above in more Taylor like form

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r(x - x_0)$$

3) $f' : \Omega \to L(\mathbb{R}^m, \mathbb{R}^n)$, so f' is function to a bigger space $\mathbb{R}^{m\times n}$ not same as $f : \Omega \to \mathbb{R}^n$.

Proposition 109. *$f'(x)$ is unique.*

Proof. Suppose two derivatives A_1,$A_2 \in L(\mathbb{R}^m, \mathbb{R}^n)$ and they satisfy the limit

$$\frac{|(A_1 - A_2)h|}{|h|} \leq \frac{\|f(x+h) - f(x) - A_1 h\|}{|h|} + \frac{\|f(x+h) - f(x) - A_2 h\|}{|h|}$$

Both terms on the right go to 0 as $h \to 0$.

Fix $\vec{v} \neq 0$, let $h = t\vec{v}$, then

$$\lim_{t\to 0} \frac{|(A_1 - A_2)tv|}{|tv|} = 0$$

hence

$$\frac{|(A_1 - A_2)v|}{|v|} = 0 \implies |(A_1 - A_2)v| \;\forall \vec{v} \implies A_1 = A_2$$

□

Lecture 12 (3/5/13)

We want to find the entry of f', choose standard base $\{e_i\}$ of $\mathbb{R}^m$, $\mathbb{R}^n$.

Suppose f is differentiable, hence

$$\lim_{h\to 0} \frac{|f(x+h) - f(x) - f'(x)h|}{|h|} = 0 \tag{3.3}$$

exists, and $f'(x) \in L(\mathbb{R}^m, \mathbb{R}^n)$. Now we take h to be specific along coordinate axes $h = te_i$ $t \in \mathbb{R}$, then

$$\lim_{t\to 0} \frac{|f(x+te_i) - f(x) - tf'(x)e_i|}{|te_i|} = 0 \,\forall\, i = 1, ..., m$$

where $f'(x)e_i$ is the ith column of the matrix $f'(x)$.

Fix $1 \le j \le n$

$$\lim_{t \to 0} \frac{f_j(x+te_i) - f_j(x) - tf'(x)_{ji}}{t} = 0 \ \ \forall i,j$$

Hence

$$f'(x)_{ji} = \lim_{t \to 0} \frac{f_j(x+te_i) - f_j(x)}{t}$$

so we have a formula for each element of matrix $f'(x)$. Clearly j gives the change in the component of $f(x)$, and i gives the change in the component of x

We will use the following notation interchangeably, $f : \mathbb{R}^m \to \mathbb{R}$, $1 \le i \le m$

$$D_i f = \partial_i f = \frac{\partial f}{\partial x_i} = f_{x_i} = \lim_{t \to 0} \frac{f(x+te_i) - f(x)}{t}$$

Remark 110. If f is differentiable at x, we denote the Jacobian matrix $J_f(x)$:

$$f'(x) = J_f(x) = \begin{pmatrix} D_1 f_1 & D_2 f_1 & \cdots & D_m f_1 \\ D_1 f_2 & & & \\ \vdots & & & \\ D_1 f_n & & \cdots & D_m f_n \end{pmatrix} \tag{3.4}$$

For special case $f : E \to \mathbb{R}$, $E \subset \mathbb{R}^m$

$$f'(x) = \nabla f(x) = (D_1 f, ..., D_m f) \tag{3.5}$$

∇ call gradient. Hence

$$\nabla f : E \to L(\mathbb{R}^m, \mathbb{R}) \sim \mathbb{R}^m$$

If ∇f is differentiable,

$$H_f(x) = f''(x) = (\nabla f)' = \begin{pmatrix} D^2_{11} f & D^2_{21} f & ... & D^2_{m1} f \\ D^2_{12} f & & & \\ \vdots & & & \\ D^2_{1m} f & & & D^2_{mm} f \end{pmatrix}$$

which is a square matrix, called Hessian. Later we will show with some mild condition, this matrix is symmetric.

Hessian is useful, e.g. one can find the maximum or minimum of $f : E \to \mathbb{R}$, $E \subset \mathbb{R}^2$.

First find critical points:

$$\nabla f = 0$$

Then apply second order approximation about critical point, say (0,0),

$$f(x, y) = f(0,0) + \nabla f(0,0) \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2}(x, y) H_f(0,0) \begin{pmatrix} x \\ y \end{pmatrix}$$

If H_f has 2 positive eigenvalue or 2 negative eigenvalue, which is a paraboloid, and then $(0,0)$ is a min or max point. If H_f has 1 positive and 1 negative eigenvalues, $(0,0)$ is a saddle point. If H_f has one of the eigenvalue being 0, then paraboloid becomes cylinder, so the test is inconclusive, then higher order Taylor expansion must be taken into account.

Note 111. The matrix $J_f(x)$ (or $\nabla f(x)$) might exist and the function might be not differentiable or not even continuous! Common examples are function that has a pitch at the origin. We'll see the situation gets worse: all directional derivative exist, the function is still not differentiable.

Example 112. find

$$\lim_{(x,y)\to(0,0)} \frac{x^2 y^2}{2x^2 + 7y^2}$$

notice: squeeze is the best option.

$$0 \le \frac{1}{2} \frac{2x^2 y^2}{2x^2 + 7y^2} \le \frac{1}{2} \frac{(2x^2 + 7y^2) y^2}{2x^2 + 7y^2} = \frac{y^2}{2} \to 0$$

One may think of writing x, y in polar, then take $r \to 0$ to ensure limit approaching from all directions. Polar is actually not quite right. It only works if

1) $f(r\cos\theta, r\sin\theta) = g(r)h(\theta)$

2) as $r \to 0$, $|h(\theta)| < M$

Example 113. Find

$$\lim_{(x,y)\to(0,0)} \frac{x^2 y}{x^2 - xy + y^2}$$

Polar

$$\frac{x^2 y}{x^2 - xy + y^2} = \frac{r^3}{r^2} \frac{\cos^2\theta \sin\theta}{1 - \cos\theta \sin\theta}$$

$$\left| \frac{\cos^2\theta \sin\theta}{1 - \cos\theta \sin\theta} \right| < \frac{1}{1 - \frac{1}{2}} = 2$$

In the following example, we try parabolic curve.

Example 114. Find

$$\lim_{(x,y)\to(0,0)} \frac{x^2 y}{x^2 - y^2}$$

try $\begin{cases} x = t + at^2 \\ y = t + bt^2 \end{cases}$, then

$$\frac{x^2 y}{x^2 - y^2} = \frac{t^3(1 + 2at + a^2t^2)(1 + bt)}{t^2[(2 + at + bt)(at - bt)]} = \frac{(1 + 2at + a^2t^2)(1 + bt)}{(2 + at + bt)(a - b)} \to \frac{1}{2(a - b)}$$

Hence the limit doesn't exist.

Theorem 115. *(Chain Rule) f, g that can be composed and g is differentiable at x and f is differentiable at $g(x)$ then $f(g(x))$ is differentiable, and*

$$[f(g(x))]' = f'(g(x)) \cdot g'(x)$$

Proof. (Same proof as in 1D, so has to be the same result just put in the correct notations.)

f is differentiable at $g(x)$ iff

$$\mu(y) = f(y + g(x)) - f(g(x)) - f'(g(x)) \cdot y \text{ and } \lim_{y\to 0} \frac{|\mu(y)|}{|y|} = 0$$

g is differentiable at x iff

$$\nu(h) = g(x + h) - g(x) - g'(x) \cdot h \text{ and } \lim_{h\to 0} \frac{|\nu(h)|}{|h|} = 0$$

then set $y = g(x + h) - g(x) = \nu(h) + g'(x) \cdot h$

$$f(g(x+h)) - f(g(x)) - f'(g(x)) \cdot (g(x+h) - g(x)) = \mu(g(x+h) - g(x))$$

or

$$f(g(x+h)) - f(g(x)) - f'(g(x)) \cdot (\nu(h) + g'(x) \cdot h) = \mu(g(x+h) - g(x))$$

If we can show the remainder

$$\lim_{h\to 0} \frac{|f'(g(x)) \cdot \nu(x) - \mu(g(x+h) - g(x))|}{|h|} = 0$$

then we complete the proof. Indeed

$$\frac{|f'(g(x)) \cdot \nu(x)|}{|h|} \leq \left\| f'(g(x)) \right\| \frac{|\nu(h)|}{|h|} \to 0$$

because $\|f'(g(x))\| < \infty$.

$$\frac{|\mu(g(x+h) - g(x))|}{|h|} = \frac{|\mu(y)|}{|y|} \frac{|\mu(g(x+h) - g(x))|}{|h|} \to 0$$

because $\frac{|\mu(y)|}{|y|} \to 0$ and

$$\frac{|\mu(g(x+h) - g(x))|}{|h|} = \frac{|\nu(h) + g'(x) \cdot h|}{|h|} \leq \frac{|\nu(h)|}{|h|} + \frac{|g'(x) \cdot h|}{|h|}$$

where $\frac{|\nu(h)|}{|h|} \to 0$ and $\frac{|g'(x) \cdot h|}{|h|} \leq \|g'(x)\| < \infty$. $\square$

Example 116.

$$[a, b] \xrightarrow{\gamma(t)} \mathbb{R}^m \xrightarrow{f} \mathbb{R}$$

$f(\gamma(t))$ is the function restricted to the curve.

Assume that γ,f are differentiable, then by chain rule

$$\frac{d}{dt} f(\gamma(t)) = f'(\gamma(t)) \cdot \gamma(t) = \nabla f(\gamma(t)) \cdot \begin{pmatrix} \gamma_1'(t) \\ \vdots \\ \gamma_m'(t) \end{pmatrix}$$

where $\nabla f(\gamma(t))$ is a row vector.

We can generalize partial derivative to directional derivative. partial derivative looks like

$$\gamma(t) = x_0 + te_i$$

$$\frac{d}{dt} f(\gamma(t)) = \nabla f(\gamma(t)) \cdot e_i$$

and directional derivative

$$\gamma(t) = x_0 + tu \quad |u| = 1$$

$$\frac{d}{dt} f(\gamma(t)) = \nabla f(\gamma(t)) \cdot u \tag{3.6}$$

Here we have assume f is differentiable, but to calculate $D_u f$, we don't need differentiability of f, using the following definition

Definition 117. Directional derivative $f : E \to \mathbb{R}$, $E \subset \mathbb{R}^m$

$$D_u f = \lim_{t \to 0} \frac{f(x+tu) - f(x)}{t} = \frac{d}{dt}\bigg|_{t=0} f(\gamma(t))$$

where $\gamma(t) = x_0 + tu \quad |u| = 1$.

Remark 118. $D_u f$ might exist $\forall\, u$, even if f is not differentiable. In this case the chain rule doesn't hold, i.e (3.6) is not true. One has to use the definition 117.

To check whether f is differentiable, one can first compute $J_f(x)$ 3.4, for $f : E \to \mathbb{R}$ uses 3.5. If $J_f(x)$ doesn't exist, then clearly f is not differentiable. If it does exist, then plug it into definition 3.3 and check the limit.

The following proposition may save some effort.

Definition 119. $f : E \to \mathbb{R}^n$, $f \in C^1(E)$ iff f is differentiable and the map

$$\begin{array}{rcl} E & \to & L(\mathbb{R}^m, \mathbb{R}^n) \\ x & \mapsto & f'(x) \end{array}$$

is continuous

Recall definition of continuous: $\forall\, \epsilon > 0$, $\forall\, x \in E$, $\exists\, \delta_{x,\epsilon} > 0$ s.t.

$$\forall\, y \in E,\ |y - x| < \delta \implies \left\| f'(x) - f'(y) \right\| < \epsilon$$

Proposition 120. *$f \in C^1(E)$ iff $D_i f_j$ is continuous $\forall\, i, j$.*

To prove this, we need old mean value theorem of 1D. However in high dimensions mean value theorem is replaced by below, we will need this to prove inverse function theorem.

Proposition 121. *(mean value theorem) $f : E \to \mathbb{R}^n$ $E \subset \mathbb{R}^m$ f is differentiable and E is convex, then $\forall\, x, y \in E$ $\exists\, z \in E$ s.t.*

$$|f(x) - f(y)| \leq \left\| f'(z) \right\| |x - y|$$

When we use the proposition, we always use it in a ball, so the convex assumption always holds.

Proof. Fix $x, y \in E$, $\forall\ t \in [0,1]$, $(1-t)x + ty \in E$, let

$$g(t) = f((1-t)x + ty)$$

then $g : E \to \mathbb{R}$, use old mean value theorem, for some t

$$|g(1) - g(0)| = |g'(t)|(1-0)$$

then by chain rule, and set $z = (1-t)x + ty$

$$|f(y) - f(x)| = |f'((1-t)x + ty) \cdot (-x + y)| \leq \left\| f'(z) \right\| |y - x|$$

$\square$

Corollary 122. *If E is convex and $\|f'(x)\| < M\ \forall x \in E$,*

$$|f(x) - f(y)| \leq M|x - y|$$

hence f is Lipschitz.

Corollary 123. *If E is convex and $f'(x) = 0\ \forall x$, then $f(x)$ is constant.*

Corollary 124. *If E is connected and $f'(x) = 0\ \forall x$, then $f(x)$ is constant.*

If the space is not connected, the statement becomes true on connected components.

Proof. Suppose the statement is false, $\exists x_1,\ x_2\ f(x_1) \neq f(x_2)$, let

$$A_1 = \{x \in E \text{ s.t. } f(x) = f(x_1)\}$$

$$A_2 = \{x \in E \text{ s.t. } f(x) \neq f(x_1)\}$$

then clearly $A_1, A_2 \neq \varnothing$ and $A_1 \cup A_2 = E$, $A_1 \cap A_2 = \varnothing$and A_1, A_2 are open. Indeed

Take $x \in A_1$, then $\exists$ a ball $B \subset E$ centered at x. Because ball is convex, $f(y) = f(x)\ y \in B$, $B \subset A_1$. $\square$

Lecture 13 (3/7/13)

We now prove proposition 120.

Proof. Suppose $f \in C^1(E)$. If $|x-y|<\delta$, $\|f'(x)-f'(y)\|<\epsilon$

Since

$$\begin{pmatrix} D_i f_1 \\ \vdots \\ D_i f_n \end{pmatrix}(x) - \begin{pmatrix} D_i f_1 \\ \vdots \\ D_i f_n \end{pmatrix}(y) = (f'(x)-f'(y)) \cdot e_i,$$

$$\begin{aligned} |D_i f_j(x) - D_i f_j(y)| &\leq |(f'(x)-f'(y))\cdot e_i| \\ &\leq \sup_{|u|\leq 1} |(f'(x)-f'(y))\cdot u| = \left\|f'(x)-f'(y)\right\| < \epsilon \end{aligned}$$

The other direction is much harder.

First show for fix j

$$f_j : E \to \mathbb{R}$$

Suppose all partial derivatives are continuous: $\exists\, \delta\ \forall\, i$

$$|D_i f_j(x) - D_i f_j(y)| < \epsilon$$

For notation convenience we write f for f_j. Fix a point x, fix ϵ, want to find δ s.t. in $B_{x,\delta}$,

$$\lim_{h\to 0} \frac{|f(x+h)-f(x)-f'(x)h|}{|h|} \to 0 \tag{3.7}$$

For any $|h|<\delta$, $h=\sum_{i=1}^m h_i e_i$, let $v_k = \sum_{i=1}^k h_i e_i$, put $v_0 = 0$. Write the telescopic sum

$$\begin{aligned} f(x+h)-f(x) &= \sum_{k=1}^m (f(x+v_k) - f(x+v_{k-1})) \\ &= \sum_{k=1}^m (f(x+v_{k-1}+h_k e_k) - f(x+v_{k-1})) \end{aligned} \tag{3.8}$$

Consider

$$\begin{aligned} g_k(t) : [0,h_k] &\to \mathbb{R} \\ t &\mapsto f(x+v_{k-1}+te_k) \end{aligned}$$

then because the partial of f exists, by mean value theorem

$$g_k(h_k) - g_k(0) = g_k'(\theta_k) h_k \quad 0 < \theta_k < h_k$$

Because $g_k'(\theta_k) = D_k f(x + v_{k-1} + \theta_k e_k)$, (3.8) becomes

$$f(x+h) - f(x) = \sum_{k=1}^{m} D_k f(x + v_{k-1} + \theta_k e_k) \cdot h_k$$

Plug into (3.7), and set $f'(x) = \nabla f(x)$

$$\begin{aligned} |f(x+h) - f(x) - f'(x)h| &= |\sum_{k=1}^{m} (D_k f(x + v_{k-1} + \theta_k e_k) - D_k f(x)) \cdot h_k| \\ &\leq \epsilon \sum_{k=1}^{m} |h_k| \leq \epsilon m |h| \end{aligned}$$

because $|(x + v_{k-1} + \theta_k e_k) - x| = |v_{k-1} + \theta_k e_k| \leq |h| < \delta$.

This shows

$$\frac{|f(x+h) - f(x) - \nabla f(x)h|}{|h|} \to 0 \tag{3.9}$$

Hence

$$\nabla f(x) \text{ is the derivative}$$

Now $f = (f_1, ..., f_n) : E \to \mathbb{R}^n$

$$\frac{|f(x+h) - f(x) - J_f(x)h|}{|h|} = \sqrt{\frac{\sum_{j=1}^{n} |f_j(x+h) - f_j(x) - \nabla f_j \cdot h|^2}{|h|^2}} \to 0$$

because each term in the sum goes to 0 by (3.9).

Hence $J_f(x)$ is the derivative, so f is differentiable.

To show $f'(x) = J_f(x)$ is continuous, use remark 106, for every entry of J_f is continuous. $\square$

3.3 Inverse Function Theorem

Recall in 1 variable, $f : E \to \mathbb{R}, E \subset \mathbb{R}$, if f is invertible, and f^{-1} is differentiable at y, then

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$$

Theorem 125. *(inverse function theorem) $f : E \to \mathbb{R}^n$, $E \subset \mathbb{R}^n$ open, f is differentiable in E and $a \in E$ s.t. $f'(a)$ is an invertible matrix and $f'(x)$ is continuous at a, then*

$$\exists\, U \subset \mathbb{R}^n \text{ open } a \in U \text{ s.t. } f : U \to f(U)$$

is invertible; $f(U)$ is open and f^{-1} is differentiable, and

$$(f^{-1})'(y) = (f'(f^{-1}(y)))^{-1}$$

and if $f \in C^1(E)$, then f^{-1} is C^1.

In differential geometry term,

Definition 126. f differentiable and invertible with differentiable inverse is called a diffeomorphism.

The theorem shows if $f'(a)$ is invertible, then f is a diffeomorphism in a neighborhood of a.

The proof relays on only one lemma: the fixed point theorem.

Definition 127. $f : X \to X$, X metric space, f is called a contraction if $\exists\, c \in \mathbb{R}$, $c < 1$ s.t. $\forall\, x, y$

$$d(f(x), f(y)) \leq cd(x, y)$$

Definition 128. x is a fixed point if $f(x) = x$.

Theorem 129. *(fixed point) $f : X \to X$*

1) If f is a contraction it has at most one fixed point.

2) If X is complete, f has exactly one fixed point.

Proof. 1) suppose $f(x_1) = x_1$ and $f(x_2) = x_2$

$$d(f(x_1), f(x_2)) = d(x_1, x_2) \leq cd(x_1, x_2) < d(x_1, x_2)$$

contradiction.

2) Pick $x_1 \in X$ any point. Let

$$x_{n+1} = f(x_n) \tag{3.10}$$

(think about this: why if we take $x_{n+1} = \frac{1}{2} f(x_n)$ or $x_{n+1} = f^2(x_n)$ would not work?)

$\{x_n\}$ is convergent, because it is Cauchy. Indeed

$$\begin{aligned}
d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) &\leq cd(x_n, x_{n-1}) \\
&= cd(f(x_{n-1}), f(x_{n-2})) \\
&\leq \\
&\vdots \\
&= c^{n-1} d(x_2, x_1)
\end{aligned}$$

so

$$\begin{aligned}
d(x_n, x_m) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_m) \\
&\leq \vdots \\
&\leq \sum_{k=n}^{m-1} d(x_k, x_{k+1}) \\
&\leq \sum_{k=n}^{m-1} c^{n-2} d(x_2, x_1) \\
&\leq \frac{1-c^m}{1-c} d(x_2, x_1) \to 0
\end{aligned}$$

for $n, m > N$, and let $N \to \infty$.

Take $n \to \infty$ to both sides of (3.10),

$$LHS \to x$$

for some $x \in X$, because X is complete.

$$\lim_{n\to\infty} RHS = f(\lim_{n\to\infty} x_n) = f(x)$$

we move limit into the function, because f is Lipschitz, i.e. continuous.

$\square$

Note 130. Looking closely to the proof above, for the theorem 2) to work, a contraction map had better be defined on D s.t. $f(D) \subset D$ domain

$$\begin{aligned}
f : [0, 2] &\to [3, 4] \\
x &\mapsto \frac{x}{2} + 3
\end{aligned}$$

is a contraction map, but no fixed point in the domain.

But proof of 1) doesn't involve recursion, so no need to have $f(D) \subset D$.

Lecture 14 (3/12/13)

Now we prove inverse function theorem. We will prove piece by piece.

1) show $\exists$ such U.

2) show $f(U)$ is open.

3) show the formula for the derivative.

Proof. 1) Fix $y \in \mathbb{R}^n$, $A = f'(a)$, let

$$\phi_y(x) = x + A^{-1} \cdot (y - f(x))$$

then ϕ_y has a fixed point iff $\exists\, x$ s.t. $y = f(x)$.

Choose U s.t. ϕ_y is a contraction of $x \in U$

$$\left\|\phi_y'(x)\right\| = \left\|1 + A^{-1} f'(x)\right\| \quad = \quad \left\|A^{-1}(A - f'(x))\right\| \leq \left\|A^{-1}\right\| \left\|A - f'(x)\right\|$$

Since f' is continuous at a, we can find a ball $B_{a,r}$ s.t.

$$\left\|A - f'(x)\right\| = \left\|f'(a) - f'(x)\right\| < \frac{1}{2\left\|A^{-1}\right\|} \tag{3.11}$$

so by mean value theorem of high dimension, $\phi_y(x)$ is a contraction on $U = B_{a,r}$, then by 1) of fixed point theorem $\exists$ at most 1 fixed point in U, so

$$f : U \quad \to \quad f(U)$$

is injective, hence f is bijective.

Here U is independent of y. But for fixed point to exist we should pick $y \in f(U)$.

2) show $y_0 \in f(U)$ is an interior point. $\exists!\, x_0 \in U$ s.t. $f(x_0) = y_0$. Make a ball $B = B_{x_0,t}$ s.t.

$$\bar{B} \subset U$$

the reason we need closure because we want to say $\bar{B}$ is closed in a complete U so $\bar{B}$ is complete.

Claim: $\exists\, s$ s.t. $\forall\, y \in B_{y_0,s}$, the contraction maps ϕ_y as defined early satisfy

$$\phi_y(B) \subset \bar{B}$$

Indeed if $|y - y_0| < s = \frac{t}{2\|A^{-1}\|}$,

$$\begin{aligned} |\phi_y(x_0) - x| &= |A^{-1}(y - f(x_0))| = |A^{-1}(y - y_0)| \\ &\leq \left\|A^{-1}\right\| |y - y_0| < \frac{t}{2} \end{aligned}$$

then

$$\begin{aligned}
|\phi_y(x) - x_0| &\leq |\phi_y(x) - \phi_y(x_0)| + |\phi_y(x_0) - x| \\
&\leq \frac{1}{2}|x - x_0| + \frac{t}{2} \\
&\leq t
\end{aligned}$$

Therefore by 2) of fixed point theorem, for each $y \in B_{y_0,s}$, $\exists ! x \in \bar{B}$ s.t. $\phi_y(x) = x$, or $y = f(x)$, so

$$B_{y_0,s} \subset f(\bar{B}) \subset f(U)$$

so $f(U)$ is open.

3) Fix $y \in f(U)$, $\exists ! x$ s.t. $f(x) = y$. Choose h s.t. $x + h \in U$, and let $y + k = f(x + h)$. Put $g = f^{-1}$

Prove g is differentiable at y.

Since $|\phi_y(x+h) - \phi_y(x)| < \frac{1}{2}|h|$ and

$$|x+h+A^{-1}(y-f(x+h))-x-A^{-1}(y-f(x))| = |h+A^{-1}(-k)| \geq \big||h| - |A^{-1}k|\big|,$$

then

$$\big||h| - |A^{-1}k|\big| < \frac{1}{2}|h| \implies \frac{1}{2}|h| < |A^{-1}k| \leq \left\|A^{-1}\right\| |k|$$

Moreover $f'(x)$ is invertible, indeed (3.11) and proof of proposition 102 give

$$\left\|A - f'(x)\right\| < \left\|A^{-1}\right\|^{-1} \implies f'(x) \text{ is invertible}$$

Now consider

$$\begin{aligned}
\frac{|g(y+k) - g(y) - f'(x)^{-1}k|}{|k|} &= \frac{|h - f'(x)^{-1}k|}{|k|} \\
&= \frac{|f'(x)^{-1}\left(f'(x)h - (f(x+h) - f(x))\right)|}{|k|} \\
&\leq \left\|f'(x)^{-1}\right\| \frac{|f(x+h) - f(x) - f'(x)h|}{|h|} 2\left\|A^{-1}\right\| \to 0
\end{aligned}$$

because if $k \to 0$, then $h \to 0$, then $\frac{|f(x+h)-f(x)-f'(x)h|}{|h|} \to 0$, therefore

$$g'(y) = f'(g(y))^{-1} \tag{3.12}$$

Lastly prove

$$\text{additionally } f \in C^1(E) \implies g \in C^1(f(U))$$

Since f' is continuous, by remark after proposition 102 f'^{-1} is continuous. g is continuous because it's differentiable, so by the formula (3.12), g' is continuous. □

Definition 131. f is open function iff $f(U)$ is open $\forall\, U$ open.

Corollary 132. *If $f : E \to \mathbb{R}^n$, $E \subset \mathbb{R}^n$ s.t. $f \in C^1(E)$, $f'(x)$ is invertible $\forall\, x \in E$, then f is open.*

Proof. By inverse function theorem, $\forall\, x\ \exists$ a ball $B_{x,r} \subset U$ s.t. $f|_{B_{x,r}}$ is invertible and $f(B_{x,r}) \subset f(U)$ open. $\forall\, y \in f(B_{x,r})$, y is an interior. □

3.4 Implicit function theorem

This is a big corollary of inverse function theorem.

Lecture 15 (3/14/13)

Suppose we want to solve the following problem: Given system

$$\begin{cases} f_1(x_1, ..., x_n; y_1, ..., y_m) = 0 \\ f_2(x_1, ..., x_n; y_1, ..., y_m) = 0 \\ \vdots \\ f_n(x_1, ..., x_n; y_1, ..., y_m) = 0 \end{cases} \tag{3.13}$$

n equations and $m+n$ variables. Solve for x_i in terms of y_j.

Use f_1, solve x_1 in terms of $x_2, ..., x_n; y_1, ..., y_m$, then plug in f_2, solve x_2 in terms of $x_3, ..., x_n; y_{1,...,}y_m$. If we are luck, we can get x_n in terms of only $y_1, ..., y_m$, then we back up to solve x_{n-1}, etc. Finally we get

$$\begin{array}{rcl} \mathbb{R}^m & \to & \mathbb{R}^n \\ (y_1, ..., y_m) & \mapsto & \begin{cases} g_1(y_1, ..., y_m) = x_1 \\ \vdots \\ g_n(y_1, ..., y_m) = x_n \end{cases} \end{array}$$

If $f_1, ..., f_n$ are linear equations, we can save a lot of time. We write (3.13) in matrix form

$$(A_x|A_y)\begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} = 0 \iff A_x\vec{x} + A_y\vec{y} = 0$$

where $(A_x|A_y)$ marries two matrices with A_x is $n \times n$ matrix, A_y is $n \times m$, $\vec{x} = (x_1, ..., x_n)$, and $\vec{y} = (y_1, ..., y_m)$. So it can be solved if A_x is invertible

$$\vec{x} = -A_x^{-1} A_y \vec{y}$$

If A_x is not invertible, i.e. some column of A_x are not linear independent, we can replace some of the columns with some columns in A_y, and rearrange

$$x_1, ..., x_n | y_1, ..., y_m$$

so that the first n variables become solvable.

Theorem 133. *(implicit function theorem)* $f : E \to \mathbb{R}^n$, $E \subset \mathbb{R}^{n+m}$, $f \in C^1(E)$ *and* $\exists\, (a,b) \in E$ *s.t.* $f(a,b) = 0$ *and* $f'(a,b) = (A_x(a,b)|A_y(a,b))$, A_n *is* $n \times n$ *invertible matrix, then*

$\exists\, V$ *open in* E *s.t.* $(a,b) \in V$ *and* $\exists\, W$ *open in* $\mathbb{R}^m$, $b \in W$, *s.t.* $\forall\, y \in W$, $\exists !\ x \in \mathbb{R}^n$ *s.t.*

$$(x,y) \in V \text{ and } f(x,y) = 0$$

this defines a function

$$\begin{array}{rcl} W & \to & \mathbb{R}^n \\ y & \mapsto & x = g(y) \end{array}$$

s.t. $f(g(y), y) \equiv 0$ *and* $g \in C^1(W)$

$$g'(y) = -A_x^{-1}(y) A_y(y)$$

Here we use derivative to approximate or linearize the function.

Paraphrase implicit function theorem in one sentence: If f is C^1, $f(a,b) = 0$ and $f'(a,b)$ has maximal rank, then locally around (a,b) we can solve for n variables and the solution is C^1.

Proof. Let

$$\begin{array}{rcl} F : E & \to & \mathbb{R}^{n+m} \\ (x,y) & \mapsto & (f(x,y), y) \end{array}$$

then $F \in C^1(E)$.

$$\begin{aligned} F(a+h, b+k) - F(a,b) &= (f(a+h, b+k) - f(a,b), k) \\ &= \left(f'(a,b) \cdot \begin{pmatrix} h \\ k \end{pmatrix} + r(h,k), k \right) \end{aligned}$$

where the remainder

$$\lim_{(h,k)\to 0} \frac{r(h,k)}{|(h,k)|} = \lim_{(h,k)\to 0} \frac{r(h,k)}{\sqrt{h^2+k^2}} = 0$$

$$\left(\begin{array}{c} f'(a,b) \cdot \left(\begin{array}{c} h \\ k \end{array} \right) \\ k \end{array} \right) = \underbrace{\left(\begin{array}{cc} A_x & A_y \\ 0 & \mathbb{I} \end{array} \right)}_{\equiv F'(a,b)} \left(\begin{array}{c} h \\ k \end{array} \right)$$

To show $F'(a,b)$ is indeed $\left(\begin{array}{cc} A_x & A_y \\ 0 & \mathbb{I} \end{array} \right)$, compute

$$\lim_{(h,k)\to 0} \frac{|F(a+h,b+k) - F(a,b) - F'(a,b) \left(\begin{array}{c} h \\ k \end{array} \right)|}{|(h,k)|} = \lim_{(h,k)\to 0} \frac{|\,(r(h,k),0)\,|}{|(h,k)|} = 0$$

Moreover clearly $\left(\begin{array}{cc} A_x & A_y \\ 0 & \mathbb{I} \end{array} \right)$ is invertible, because

$$\left(\begin{array}{cc} A_x & A_y \\ 0 & \mathbb{I} \end{array} \right) \left(\begin{array}{c} h \\ k \end{array} \right) = \left(\begin{array}{c} 0 \\ 0 \end{array} \right) \implies k = 0 \implies h = 0$$

hence injective, so bijective.

Use inverse function theorem, $\exists\, V, U$, $V \subset E$ open and $(a,b) \in V$, $U \subset \mathbb{R}^{n+m}$ open, $U = F(V)$,

$$F : V \to U$$

is invertible, and the inverse is $C^1(U)$.

Notice $W \subset \mathbb{R}^m$, $y \in W$ means $(0,y) \in U$. Prove y is interior. $\forall\, \delta$, since U is open, $\exists\, (u,v) \in U$ s.t.

$$|(u,v) - (0,y)| < \delta \implies |v - y| < \delta \implies v \in W$$

so W is open.

$\forall\, y \in W$, $(0,y) \in U$, because F is invertible, $\exists !\ x$ s.t.

$$(x,y) \in V \text{ and } F(x,y) = (0,y) \implies f(x,y) = 0$$

This defines

$$\begin{array}{rcl} g : W & \to & \mathbb{R}^n \\ y & \mapsto & x \end{array}$$

Prove that $g \in C^1(W)$.

$$F(g(y), y) = (0, y) \iff (g(y), y) = F^{-1}(0, y) \tag{3.14}$$

i.e.

$$\begin{array}{ccccc} W & \to & \mathbb{R}^{n+m} & \to & \mathbb{R}^{n+m} \\ y & \mapsto & (0, y) & \mapsto & F^{-1}(0, y) \end{array}$$

Differentiate RHS of (3.14) with respect to y,

$$\begin{pmatrix} g'(y) \\ \mathbb{I} \end{pmatrix} = (F^{-1})' \cdot \begin{pmatrix} 0 \\ \mathbb{I} \end{pmatrix}$$

Since F is invertible, by inverse function theorem $\forall\, y \in W$, $(F^{-1})'(0, y) = (F'(g(y), y))^{-1}$.

Thus

$$F'(g(y), y) \begin{pmatrix} g'(y) \\ \mathbb{I} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbb{I} \end{pmatrix}$$

That is

$$\begin{pmatrix} A_x & A_y \\ 0 & \mathbb{I} \end{pmatrix} \begin{pmatrix} g'(y) \\ \mathbb{I} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbb{I} \end{pmatrix} \implies A_x g'(y) + A_y = 0$$

Therefore

$$g'(y) = -A_x' A_y$$

$\square$

To better understand why implicit function theorem works, we should look at

Theorem 134. *(constant rank theorem) $f : E \to \mathbb{R}^n$ $E \subset \mathbb{R}^m$ open, $f \in C^1(E)$. $\forall\, x \in E$ the rank of $f'(x)$ is $k \leq m, n$ (k is independent of x).*

$$\begin{array}{ccc} \mathbb{R}^m & \xrightarrow{f'} & \mathbb{R}^k \\ & & \uparrow p \\ E & \xrightarrow{f} & \mathbb{R}^{m+n} \end{array}$$

P is identity on $\mathbb{R}^k$. Fix a point $a \in E$, $A = f'(a)$, then $\exists U \subset E$ open, $a \in U$ and $V \subset \mathbb{R}^m$ and

$$U \xrightarrow{G} V$$

G is a diffeomorphism s.t.

$$\begin{array}{ccc} & V & \xrightarrow{A} & A(V) & \\ G\uparrow & \downarrow H & & P\uparrow & \downarrow S \\ & U & \xrightarrow{f} & \mathbb{R}^{n+m} & \end{array}$$

commute. S section, one side inverse, $PS = I$.

Rudin doesn't use the diagram, instead he says

$$\forall\, x \in V \quad f(H(x)) = Ax + \phi(Ax)$$

where $\phi : A(V) \to \ker P$.

According to the diagram,

$$f(H(x)) = S(Ax)$$

The two statements are the same, i.e.

$$S(Ax) = Ax + \phi(Ax)$$

This theorem says locally differentiation is a good approximation of the function. That is what used in implicit function thereon.

3.5 More about Derivatives

Lecture 16 (3/26/13)

Theorem 135. *(Clairaut) $f(x,y) : E \to \mathbb{R}$, $E \subset \mathbb{R}^2$ open, $D_1 f$ and $D_2 f$ exists $\forall\,(x,y) \in E$ $D_{21} f$ exists $\exists\,(\bar{x},\bar{y}) \in E$ s.t. $D_{21} f$ is continuous at $(\bar{x},\bar{y})$. Then $D_{12} f$ also exists at $(\bar{x},\bar{y})$ and*

$$D_{21} f(\bar{x},\bar{y}) = D_{12} f(\bar{x},\bar{y})$$

We need 2D version of mean value theorem to prove it.

Lemma 136. *f on $E \subset \mathbb{R}^2$, E open and both $D_1 f$, $D_2 f$ exist everywhere. Assume that $\exists$ rectangle $[x, x+a] \times [y, y+b] \subset E$, then $\forall\, x, y \in [x, x+a] \times [y, y+b]$, $\exists\,(\bar{x},\bar{y}) \in (x, x+a) \times (y, y+b)$ s.t.*

$$f(x+a, y-b) - f(x+a, y) - f(x, y+b) + f(x,y) = abD_{21} f(\bar{x},\bar{y})$$

Proof. Assume $a, b > 0$, apply mean value to $f(t, y+b) - f(t, y)$ $t \in [x, x+a]$

$$f(x+a, y+b) - f(x+a, y) - (f(x, y+b) - f(x, y)) = a(D_1 f(\bar{x}, y+b) - D_2 f(x, \bar{y}))$$

for some $\bar{x} \in (x, x+a)$. Then apply mean value theorem on y

$$a(D_1 f(\bar{x}, y+b) - D_2 f(x, \bar{y})) = ab D_{21} f(\bar{x}, \bar{y})$$

some $\bar{y} \in (y, y+b)$. □

Prove Clairaut

Proof. $(\bar{x}, \bar{y}) \in E$ where $D_{21} f$ is continuous $\forall\, \epsilon > 0\ \exists\, \delta_\epsilon > 0$ s.t.

$$|D_{21} f(u + \bar{x}, v + \bar{y}) - D_{21} f(\bar{x}, \bar{y})| < \epsilon$$

if $d((u + \bar{x}, v + \bar{y}), (\bar{x}, \bar{y})) < \delta_\epsilon$, where u, v are the lengths of sides of rectangle contained in $B_{(\bar{x}, \bar{y}), \delta_\epsilon}$.

By the lemma

$$\underbrace{f(\bar{x}+v, \bar{y}+u) - f(\bar{x}+v, \bar{y}) - f(\bar{x}, \bar{y}+u) + f(\bar{x}, \bar{y})}_{\equiv \Delta f} = uv D_{21} f(\tilde{x}, \tilde{y})$$

where $(\tilde{x}, \tilde{y}) \in u, v$ rectangle $\implies (\tilde{x}, \tilde{y}) \in B_{(\bar{x}, \bar{y}), \delta_\epsilon}$. Then

$$\left| \frac{\Delta f}{uv} - D_{21} f(\bar{x}, \bar{y}) \right| \leq \left| \frac{\Delta f}{uv} - D_{21} f(\tilde{x}, \tilde{y}) \right| + |D_{21} f(\tilde{x}, \tilde{y}) - D_{21} f(\bar{x}, \bar{y})| < \epsilon$$

Take $u \to 0$, get

$$\left| \frac{D_2 f(\bar{x}+v, \bar{y}) - D_2 f(\bar{x}, \bar{y})}{v} - D_{21} f(\bar{x}, \bar{y}) \right| < \epsilon$$

showing

$$\lim_{v \to 0} \frac{D_2 f(\bar{x}+v, \bar{y}) - D_2 f(\bar{x}, \bar{y})}{v}$$

exists and

$$D_{12} f(\bar{x}, \bar{y}) = D_{21} f(\bar{x}, \bar{y})$$

□

Theorem 137.
$$f(t) = \int_a^b \phi(x,t) d\alpha(x)$$
$\forall\, c \le t \le d$, $\phi(\cdot, t) \in \mathcal{R}(\alpha)$, $D_2\phi(x,t)$ exists, and $\forall\, \epsilon$, $\exists\, \delta \forall\, x \in [a,b]$,
$$|s-t| < \delta \implies |D_2\phi(x,s) - D_2\phi(x,t)| < \epsilon$$
then
$$f'(t) = \int_a^b D_2\phi(x,y) d\alpha(x)$$
$\forall\, c < t < d$.

Proof. Fix a point $c < t < d$,
$$\frac{f(t+h) - f(t)}{h} = \int_a^b \frac{\phi(x,t+h) - \phi(x,t)}{h} d\alpha(x)$$
Since $D_2\phi(x,t)$ exists everywhere, by mean value theorem
$$RHS = \int_a^b \frac{h}{h}\phi(x,u(h)) d\alpha(x)$$
where $u(h) \in (t, t+h)$ and $u(h) \to t$ as $h \to 0$ for all x.

Take any sequence $u_n \to t$, since $D_2\phi(x,u_n)$ converges uniformly in x, indeed $\forall\, \epsilon$, $\exists\, \delta\ \forall\, x \in [a,b]$,
$$|s-t| < \delta \implies |D_2\phi(x,s) - D_2\phi(x,t)| < \epsilon$$
By the sequential convergence of a function, i.e. $\lim_{p \to q} f(p) = L \iff \forall$ sequence $p_n \to p$, $\lim_{n\to\infty} f(p_n) = L$,
$$D_2\phi(x,u(h)) \to D_2\phi(x,t)$$
as $h \to 0$. □

4 Integral Theory

We'll do integral theory first, because it is kind of easier than measure theory, although integral theory is more general than measure theory, and it doesn't depend on the specific type of measures.

Even without knowing the specific measure, we can do a lot things: monotone convergence, Futon's lemma, dominate convergence, etc.

Reference: Rudin real and complex analysis chapter 1 for integral theory. Ash probability & measure theory chapters 1, 2 for construction of measure.

4.1 Measurability

The idea of measurable set is from integration.

$$f : E \to \mathbb{R}/\mathbb{C}$$

$E \subset \mathbb{R}^n$, don't have to assume $f \geq 0$.

Recall in Riemann integral, domain of integral is partitioned. Lebesgue cuts codomain and find preimage

$$f^{-1}(\underbrace{[y_i, y_{i+1}]}_{E_i})$$

and assign it a size, called measure

$$\mu(f^{-1}(E_i))$$

and we can replace f by step function that is underneath f, and then the value of the integral is approximated by

$$\left(\sum_i^n \inf_{E_i} f\right) \left(\mu(f^{-1}(E_i))\right) \tag{4.1}$$

If we can make E_i finer and finer, we get integral.

The big problem is how to define measure in $\mathbb{R}^n$.

We want it to have several properties:

0) $\mu : P(\mathbb{R}^n) \to [0, \infty]$, $P(\mathbb{R}^n)$ is the power set

1) $\mu(A \cup B) = \mu(A) + \mu(B)$ if $A \cap B = \varnothing$.

This may not be enough, because in the process of (4.1), we need to take limit. So better property is desired

2) $\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ if $A_i \cap A_j = \varnothing \ \forall\, i \neq j$.

The infinite sum is allowed to ∞

3) $\mu(\varnothing) = 0$

We want our measure to have ordinary meanings i.e. measure of interval gets the length, measure of cube gets the volume

4) $\mu(n\text{-cell}) = \text{volume} = \text{product of the edges}$.

5) $A \subset \mathbb{R}^n$ $v \in \mathbb{R}^n$, $A + v = \{a + v, a \in A\}$, then

$$\mu(A) = \mu(A + v)$$

$\forall\, v$. called translational invariant.

This would be a Lebesgue measure, but turn out that such μ doesn't exist. We have to drop either

$$2) + 5)$$

or

$$0)$$

we'll drop 0). So we are working in a smaller subset of $P(\mathbb{R}^n)$.

Note: later we will do abstract measure then 0)-5) are fine.

Definition 138. X a set $\mathcal{M}$ a collection of subsets of X s.t.

1) $A_n \in \mathcal{M} \implies \cup_{n=1}^{\infty} A_n \in \mathcal{M}$

2) $X \in \mathcal{M}$

3) $A \in \mathcal{M} \implies A^c \in \mathcal{M}$

$(X, \mathcal{M})$ is called a measure space, and $A \in \mathcal{M}$ are called measurable. $\mathcal{M}$ is called a σ algebra or σ field.

Recall If 1) is replaced by $A_n \in \mathcal{M} \implies \cup_{n=1}^{N} A_n \in \mathcal{M}$ finite unions, $\mathcal{M}$ is an algebra or a field.

Remark 139. This automatically implies $\varnothing \in \mathcal{M}$, finite union $\in \mathcal{M}$, finite or infinite intersection $\in \mathcal{M}$, all difference of 2 sets in $\mathcal{M}$, and $P(X) \in \mathcal{M}$

Note 140. The definition above is very similar to the definition of a topology, with one subtle difference. In 1) we used countable union. Recall in definition of Topology no requirement on countable union.

Lecture 17 (3/28/13)

Definition 141. $(X, \mathcal{M})$ is a measure space and $(Y, \mathcal{F}_0)$ is a topology and

$$f : X \to Y$$

is measurable iff

$$f^{-1}(V) \in \mathcal{M} \quad \forall\, V \in \mathcal{F}_0$$

That is f is measurable iff preimage of open is measure.

Remark 142.

$$X \xrightarrow{f} Y \xrightarrow{g} Z$$

X is measure space, Y,Z are topology, f is measurable, g is continuous, then $g \circ f$ is measurable.

Question 143. *Is continuous function measurable?* $\iff$ *Is open set measurable?*

Proposition 144. *X a set $\mathcal{F}_0 \subset \mathcal{P}(X)$ $\exists\, m$ a σ algebra on X which is the smallest σ algebra containing $\mathcal{F}_0$.*

Proof. Let $a \in A$, so that m_a is a σ algebra containing $\mathcal{F}_0$, put

$$m = \bigcap_{a \in A} m_a \supset \mathcal{F}_0$$

$A \neq \varnothing$ because $\mathcal{P}(X)$ is a σ algebra. Want to prove that m is σ algebra.

If $B \in m \implies B \in m_a\ \forall\, a \in A \implies B^c \in m_a,\ \forall\, a \in A \implies B^c \in m$.

Similar verify other properties. $\square$

Use this idea we will construct the small σ algebra of all open sets.

Remark 145. $(X, \mathcal{F}_0)$ a topology $\exists\, m$ the smallest σ algebra s.t.

$$\mathcal{F}_0 \subset m$$

we denote $m := B(X)$ the Borel sets.

e.g. countable intersection of open sets is not open but it is in Borel.

Example 146. $B(\mathbb{R})$ contains (a, b), $[a, b]$, and

$$[a, b) = \bigcap_{n \in \mathbb{N}} (a_n, b) \quad a_n \uparrow a$$

and others

Exercise 147. find non Borel sets in $\mathbb{R}$.

Remark 148.

$$f : (X, B(X)) \to (Y, \mathcal{F}_0)$$

$(Y, \mathcal{F}_0)$ topology, then

$$f \text{ is continuous} \implies f \text{ is measurable}$$

Later we will show measurable functions are quite stable. Sum, product, of measurable functions are measurable. limit of sequence of measurable function are measurable.

Proposition 149. *$F : \mathbb{R}^2 \to Y$, Y a top, F is continuous X measure space f, g are measurable functions*

$$\begin{array}{ccccc} X & \xrightarrow{(f,g)} & \mathbb{R}^2 & \xrightarrow{F} & Y \\ x & \mapsto & (f(x), g(x)) & \mapsto & F(f(x), g(x)) \end{array}$$

then $F(f, g)$ is measurable.

Proof. use previous remark, we can show

$$(f, g) : X \to \mathbb{R}^2$$

is measurable. If $V \subset \mathbb{R}^2$ open show $(f, g)^{-1}(V)$ is measurable

Recall in analysis we showed open in $\mathbb{R}^2$ is equal to countable union of open rectangles, i.e.

$$V = \bigcup_{n \in \mathbb{N}} (a_n, b_n) \times (c_n, d_n)$$

so

$$(f, g)^{-1}(V) = \bigcup_{n \in \mathbb{N}} (f, g)^{-1}(a_n, b_n) \times (c_n, d_n)$$

Since

$$\begin{aligned} (f, g)^{-1}(a_n, b_n) \times (c_n, d_n) &= \{x \in X \text{ s.t. } f(x) \in (a_n, b_n) \text{ and } g(x) \in (c_n, d_n)\} \\ &= f^{-1}(a_n, b_n) \cap g^{-1}(c_n, d_n) \end{aligned}$$

this shows it is measurable. □

Corollary 150. *f, g measurable $X \to \mathbb{R}$, $f + g$, $f \cdot g$ are measurable.*

Proof. take $F(x, y) = x + y$ or $x \cdot y$ □

Proposition 151. *$f : X \to \mathbb{R}$ is measurable iff $f^{-1}(a, \infty)$ is measurable $\forall\, a \in \mathbb{R}$.*

Proof. assume $\forall\, a \in \mathbb{R}$ $f^{-1}(a + \infty)$ is measurable. Prove that $f^{-1}(V)$ is measurable $\forall\, V \subset R$ open.

Recall open set in $\mathbb{R}$ can be written as disjoint union of open intervals

$$V \text{ is open} \iff V = \coprod_{n \in \mathbb{N}} (a_n, b_n)$$

So our task becomes to prove $f^{-1}(a_n, b_n)$ is measurable.

$$f^{-1}(a_n, b_n) = f^{-1}(a_n, \infty) \cap f^{-1}(-\infty, b_n)$$

$f^{-1}(-\infty, b_n) = \left(f^{-1}[b_n, \infty)\right)^c$, now prove $f^{-1}[b_n, \infty)$ is measurable, because

$$[b_n, \infty) = \bigcap_{k=1}^{\infty} (p_k, \infty)$$

where $p_k \uparrow b$. $\square$

Remark 152. In the proportion above one can replace (a, ∞) with $[a, \infty)$ or $(-\infty, a)$, or $(-\infty, a]$.

Recall limit of Riemann integral function is not necessary Riemann integrable, we need uniform convergence for Riemann, but for Lebesgue only pointwise is enough. It all starts below.

Proposition 153. *f_n a sequence of measurable functions*

1)

$$\sup_{n \in \mathbb{N}} f_n(x) = g(x)$$

is measurable, including the case $g(x)$ goes to ∞. $\inf_{n \in \mathbb{N}} f_n(x) = h(x)$ is measurable.

2)

$$\limsup_{n \to \infty} f_n \quad \liminf_{n \to \infty} f_n$$

are measurable.

3) If f_n is convergent,

$$\lim_{n \to \infty} f_n$$

is measurable.

Proof. 1) show $g^{-1}(a, \infty)$ is measurable $\forall\, a \in \mathbb{R}$,

$$g^{-1}(a, \infty) = \bigcup_{n \in \mathbb{N}} f_n^{-1}(a, \infty)$$

Indeed if $x \in f_n^{-1}(a, \infty) \iff f_n(x) > a$

$$g(x) \geq f_n(x) > a \;\forall\, n \in \mathbb{N} \implies g^{-1}(a, \infty) \supset \bigcup f_n^{-1}(a, \infty)$$

Conversely, fix $x \in g^{-1}(a,\infty) \implies g(x) > a$. $\exists\, n \in \mathbb{N}$ s.t.

$$g(x) \geq f_n(x) > a \implies x \in f_n^{-1}(a,\infty)$$

so

$$g^{-1}(a,\infty) \subset \bigcup f_n^{-1}(a,\infty)$$

2)

$$\limsup_{n\to\infty} f_n(x) = \lim_{n\to\infty}\left(\sup_{m\geq n} f_m(x)\right) = \inf_{n\geq 1}\sup_{m\geq n} f_m(x)$$

$\sup_{m\geq n} f_m(x)$ is measurable for each n, so

$$\inf_{n\geq 1}\sup_{m\geq n} f_m(x)$$

is measurable by 1).

3) = 1) + 2) $\square$

Remark 154. $\max(f,g)$, $\min(f,g)$ are measurable if f,g are measurable. Proof write

$$\max(f,g) = \frac{(f+g)+|f-g|}{2}$$

and $|\ |$ is continuous function.

Remark 155. Put

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

$f^+ = \max(f,0)$, $f^- = -\min(f,0)$ are measurable.

4.2 Lebesgue Integral

We'll define Lebesgue integral on positive functions

Remark 156. $E \subset X$ measurable space . Characteristic function

$$\chi_E = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}$$

$$\chi_E \text{ is measurable} \iff E \text{ is measurable}$$

Proof. V open in $\mathbb{R}$. If $1 \in V$, $0 \notin V$, then $\chi^{-1}(V) = E$. If $1, 0 \in V$, $\chi^{-1}(V) = X$. If $0 \in V$, $1 \notin V$, then $\chi^{-1}(V) = E^c$. $\square$

This shows if there is non-measurable set, there is non-measurable functions. For finding non Lebesgue measurable set in $\mathbb{R}$, one has to have axiom of choice.

Definition 157. X measurable space.

$$s(x) = \sum_{i=1}^{n} \chi_{E_i}(x) c_i$$

$c_i \in \mathbb{R}$, $E_i \subset X$. s is called a simple function.

s is measurable iff E_i is measurable $\forall\, i = 1, ..., n$.

We will use simple to approximate f. If f is bounded, we'll show the approximation is uniform, then we will use them to evaluate the integral.

Lecture 18 (4/2/13)

Proposition 158. *$f : X \to \mathbb{R}$, $f \geq 0$ $\exists$ s_n a sequence of simple functions s.t.*

1) $s_n \leq f$

2) $s_n \leq s_{n+1}$

3) if f is measurable, s_n is measurable

4) $s_n \to f$ pointwise and if f is bounded from above, it converges uniformly.

Proof. The proof is constructive, we will actually make s_n. $0 \leq i \leq n2^n$

$$E_i = f^{-1}([i2^{-n}, (i+1)2^{-n}])$$

f is measurable $\implies$ E_i is measurable. Let $k_n = f^{-1}[n, \infty)$, put

$$s_n = \sum_{i=0}^{n2^n - 1} \frac{i}{2^n} \chi_{E_i} + n\chi_{k_n}$$

showing 1)-3) are true.

4) Fix x in the domain, find n s.t. $f(x) < n$ so $\exists\, i$ s.t.

$$\frac{i}{2^n} \leq f(x) < \frac{i+1}{2^n}$$

for each $x \in E_i$

$$0 \leq f(x) - s_n(x) = f(x) - \frac{i}{2^n} < \frac{i+1}{2^n} - \frac{i}{2^n} \to 0$$

so here n depends on x, but if $f(x) \leq M$ $x \in D(f)$, fix n s.t. $n \geq M$ then repeat the procedure, get uniform convergence. $\square$

Before we get to Lebesgue integral, let us define measure.

Definition 159. $(X, \mathcal{M})$ measure space

$$\begin{array}{rcl} \mu : \mathcal{M} & \to & [0, \infty] \\ E & \mapsto & \mu(E) \end{array}$$

$E_n \in \mathcal{M}$ is a countable collection of sets and

1) $E_n \cap E_m = \varnothing$ if $n \neq m$

2) $\mu(\cup_{n\in\mathbb{N}} E_n) = \sum_{n=1}^{\infty} \mu(E_n)$

3) $\exists\, A \in \mathcal{M}$ s.t. $\mu(A) < \infty$.

call μ a measure.

Remark 160. If $\mu(X) = 1$, we call it a probability measure.

Proposition 161. *1)* $\mu(\varnothing) = 0$

2) $\mu(\cup_{n=1}^{N} E_n) = \sum_{n=1}^{N} \mu(E_n)$ *for* $E_i \cap E_j = \varnothing$, $i \neq j$

3) $A \subset B \implies \mu(A) \leq \mu(B)$

4) $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$

5) if $E_1 \subset E_2 \subset ... \subset E_n \subset E_{n+1} \subset ...$, *then*

$$\mu(E_n) \to \mu(\bigcup_{n\in\mathbb{N}} E_n)$$

called continuity from below

6) $\exists\, i$ *s.t.* $\mu(E_i) < \infty$, $E_1 \supset E_2 \supset ... \supset E_n \supset D_{n+1} \supset ...$, *then*

$$\mu(E_n) \to \mu(\bigcap_{n\in\mathbb{N}} E_n)$$

called continuity from above.

Proof. 1) A,$\varnothing, ..., \varnothing$ they are disjoint

$$A \cup (\bigcup \varnothing) = A \text{ and } \mu(A) < \infty$$

$$\mu(A) + \sum \mu(\varnothing) = \mu(A) \implies \mu(\varnothing) = 0$$

2) take E_1, $E_2, ...,$ E_N, $\varnothing$, $\varnothing, ...$

3) $B = A \coprod (B - A)$ disjoint union

$$\mu(B) = \mu(A) + \mu(B - A) \implies \mu(B) \geq \mu(A)$$

4)

$$A \cup B = (A - B) \coprod (A \cap B) \coprod (B - A)$$

$$A = (A - B) \coprod A \cap B \qquad B = (B - A) \coprod A \cap B$$

then use 2)

5) make $B_1 = E_1$, $B_2 = E_2 - E_1$,..., $B_n = E_n - E_{n-1}$

$$B_i \cap B_j = \varnothing \quad i \neq j$$

$$\coprod_{i=1}^{n} B_i = E_n$$

$$\mu(E_n) = \sum_{i=1}^{n} \mu(B_i)$$

$$\mu(E_n) \to \sum_{i=1}^{\infty} \mu(B_i) = \mu(\coprod_{i=1}^{\infty} B_i) = \mu(\bigcup_{n \in \mathbb{N}} E_n)$$

6) Assume $\mu(E_1) < \infty$ take complement wrt E_1

$$E_1 \supset E_2 \supset ...$$

because

$$\varnothing \subset E_1 - E_2 \subset E_1 - E_3 \subset ...$$

by 5)

$$\mu(E_1 - E_n) \to \mu(\bigcup_{n \in \mathbb{N}} (E_1 - E_n))$$

so

$$\mu(E_1) - \mu(E_n) \to \mu(E_1 - \bigcap E_n)$$

so

$$\mu(E_n) \to \mu$$

If $\mu(E_1) = \infty$, start out everything from $\mu(E_i)$. $\square$

Now we do integration. We don't have to integrate the whole domain.

Definition 162. $(X, \mathcal{M})$ measure space, $f \geq 0$ a measurable function $0 \leq s \leq f$ a simple measure function, $E \subset X$ measurable and $\mu : m \to [0, \infty]$ a measure

$$\int_E f d\mu = \sup_{0 \leq s \leq f} \int_E s d\mu$$

here $s = \sum_{i=1}^n c_i \chi_{E_i}$, we define

$$\int_E s d\mu = \sum_{i=1}^n c_i \mu(E_i \cap E)$$

Remark 163. Why do we need f to be measurable? Clearly as long as s is measurable things are fine. But when we prove linearity for integration, we do need f to be measurable. However f is measurable is not a big burden. E.g. Dirichlet function is not Riemann but is Lebesgue, and Dirichlet is measurable because the set of rationals is Borel set.

Proposition 164. *1)* $f \leq g \implies \int_E f d\mu \leq \int_E g d\mu$

2) $A \subset B \implies \int_A f d\mu \leq \int_B f d\mu$

3) $\mu(E) = 0 \implies \int_E f d\mu = 0$ *even if* f *is divergent*

4) $f \equiv 0 \implies \int_E f d\mu = 0$ *even if* $\mu(E) = \infty$

Proof. 3) s_n is bounded then take $\mu(E_i \cap E) = 0$

4) assume $(X, \mathcal{M})$ is σ finite i.e. $\forall\, E \in \mathcal{M}$

$$E = \coprod_{i=1}^{\infty} E_i \text{ s.t. } \mu(E_i) < \infty\, \forall\, i$$

then the only possible $s_n \equiv 0$ □

Proposition 165. *if* s *is measurable simple function* ≥ 0,

$$\begin{aligned} \phi : m &\to [0, \infty] \\ E &\mapsto \int_E s d\mu \end{aligned}$$

then ϕ *is a measure.*

Lecture 19 (4/4/13)

Proof. $s = \sum_{i=1}^M c_i \chi_{E_i}$, $E_i \in \mathcal{M}$, let $A_1, A_2 ..., A_n, ...$ $A_n \cap A_m = \varnothing$ if

$n \neq m$ $A_n \in \mathcal{M}$,

$$\begin{aligned}
\phi(\coprod_{n=1}^{\infty} A_n) &= \int_{\coprod_{n=1}^{\infty} A_n} \sum_{i=1}^{M} c_i \chi_{E_i} d\mu \\
&= \sum_{i=1}^{M} c_i \mu(\coprod_{n=1}^{\infty} E_i \cap A_n) \because \text{finite sum can swap} \\
&= \sum_{i=1}^{M} c_i \sum_{n=1}^{\infty} \mu(E_i \cap A_n) = \sum_{n=1}^{\infty} \sum_{i=1}^{M} c_i \mu(E_i \cap A_n) \\
&= \sum_{n=1}^{\infty} \int_{A_n} c_i d\mu = \sum_{n=1}^{\infty} \phi(A_n)
\end{aligned}$$

□

Remark 166. After we have the theorem that allows us to take limit, then we can go back to prove this for measurable function f. Later we will show since ϕ is a measure, we can do

$$\int_E g d\phi = \int_E g f d\mu$$

or write $d\phi = f d\mu$. If given ϕ,μ can one find f? (i.e. $f = \frac{d\phi}{d\mu}$) proof requires monotone convergence theorem.

We now prove linearity of integration, multiplied by a constant is trivial.

Proposition 167. *s, t simple measurable, $s, t \geq 0$*

$$\int_X (s+t) d\mu = \int_X s d\mu + \int_X t d\mu$$

Proof. $s = \sum_{i=1}^{N} c_i \chi_{A_i}$,$t = \sum_{i=1}^{M} c_i \chi_{B_i}$, assume

$$X = \coprod_{i=1}^{N} A_i = \coprod_{i}^{M} B_i$$

This is doable, just put $A_i = s^{-1}(\{c_i\})$. Since each $x \in A_i \cap B_j$

$$X = \coprod_{i,j} A_i \cap B_j \text{ and } (s+t)(x) = c_i + d_j$$

$$\int_{A_i \cap B_j} (s+t) d\mu = (c_i + d_j)\mu(A_i \cap B_j) = \int_{A_i \cap B_j} s d\mu + \int_{A_i \cap B_j} t d\mu$$

$$\begin{aligned}\int_{\coprod_{i,j} A_i\cap B_j}(s+t)d\mu &= \sum_{i,j}^{N,M}\int_{A_i\cap B_j}(s+t)d\mu \\ &= \sum \int_{A_i\cap B_j} sd\mu + \int_{A_i\cap B_j} td\mu = \int_X sd\mu + \int_X td\mu\end{aligned}$$

□

4.3 Monotone Convergence

This is our first limit convergent theorem, very fundamental, used to prove everything henceforth.

Theorem 168. *(Lebesgue monotone convergence)* $(X, \mathcal{M}, \mu)$ $f_n \geq 0$ *measurable* $f_n \leq f_{n+1}$ *for each* x, $f_n(x) \to f(x)$ *pointwise then*

$$\lim_{n\to\infty}\int_X f_n d\mu = \int_X f d\mu$$

Proof. We allow $f_n(x)$ to diverge somewhere. $f_n \leq f_{n+1} \implies \int_X f_n d\mu \leq \int_X f_{n+1} d\mu$ hence $\int_X f_n d\mu$ is an increasing sequence, so

$$\int_X f_n d\mu \to \alpha \in [0,\infty]$$

$$f_n \leq f \implies \int_X f_n d\mu \leq \int_X f d\mu$$

and

$$\alpha = \sup \int_X f_n d\mu \implies \alpha \leq \int_X f d\mu$$

Now the other way.

Fix $s \geq 0$ a measurable simple function s.t. $s \leq f$ Fix $c \in \mathbb{R}$ $0 < c < 1$

$$\begin{aligned} X_n &= \{x \in X \text{ s.t. } f_n(x) \geq cs(x)\} \\ &= (f_n - cs)^{-1}[0,\infty]\end{aligned}$$

so X_n measurable.

$$\alpha \leftarrow \int_X f_n d\mu \geq \int_{X_n} f_n d\mu \geq c\int_{X_n} sd\mu$$

$$\phi(x) = \int_{X_n} sd\mu$$

is a measure

$$f_n(x) \geq cs(x) \implies f_{n+1}(x) \geq f_n(x) \geq cs(x)$$

so $X_n \subset X_{n+1}$, then

$$\lim_{n\to\infty} \int_{X_n} sd\mu = \int_{\cup X_n} sd\mu = \int_X sd\mu$$

Given $x \in X$, since $cs(x) < s(x) \leq f(x)$, and $f(x) = \sup_n f_n(x)$, we can find N_x s.t. $x \in X_{N_x} \implies X = \cup X_n$, so

$$\int_{\cup X_n} sd\mu = \int_X sd\mu$$

Then$\lim_{c\to 1^-}(\alpha \geq c\int_X sd\mu)$ gives

$$\sup_{0\leq s\leq f}(\alpha \geq \int_X sd\mu) \quad \forall\, 0 \leq s \leq f \text{ measurable}$$

so $\alpha \geq f_X f d\mu$. $\square$

Proposition 169. $(X, \mathcal{M}, \mu)$, $f_n \geq 0$ *measurable*

$$\int_X (\sum_{n=1}^{\infty} f_n) d\mu = \sum_{n=1}^{\infty} \int_X f_n d\mu$$

Proof. First consider $\int_X (f_1 + f_2) d\mu$, take

$$s_N^i \nearrow f_i \quad 0 \leq s_N^i \leq f_i$$

$i = 1, 2$, s_N^i simple measurable then

$$(s_N^1 + s_N^2) \nearrow (f_1 + f_2)$$

by monotone convergence

$$\int_X (s_n^1 + s_n^2) d\mu \to \int_X (f_1 + f_2) d\mu$$

hence

$$\int_X s_n^1 d\mu + \int_X s_n^2 d\mu \to \int_X f_1 d\mu + \int_X f_2 d\mu$$

$\sum_{n=1}^{N} f_n$ is increasing because $f_n \geq 0$, by induction

$$\sum_{n}^{N} \int_X f_n d\mu = \int_X \sum_{n}^{N} f_n d\mu$$

then use monotone convergence again

$$\lim_{N\to\infty} \sum_{n}^{N} \int_X f_n d\mu \quad = \quad \lim_{N\to\infty} \int_X \sum_{n}^{N} f_n d\mu = \int_X \sum_{n}^{\infty} f_n d\mu$$

$\square$

Example 170. counting measure $(\mathbb{N}, P(\mathbb{N}), \mu)$

$$\mu(A) = \begin{cases} \#A & \text{if } A \text{ finite} \\ \infty & \text{if } A \text{is not finite} \end{cases}$$

Then

$$\begin{aligned} \int_{\mathbb{N}} f d\mu &= \int_{\coprod_{n\in\mathbb{N}}\{n\}} f d\mu = \sum_{n=1}^{\infty} \int_{\{n\}} f d\mu \\ &= \sum_{n=1}^{\infty} f(n)\mu(\{n\}) = \sum_{n=1}^{\infty} f(n) \end{aligned}$$

We can now use counting measure to show swapping summation, let $f_n : \mathbb{N} \to [0, \infty)$ and

$$\begin{aligned} f : \mathbb{N}^2 &\to [0, \infty) \\ n, m &\mapsto f_{nm} = \sum_{n=1}^{m} f_n \end{aligned}$$

$$\int_{\mathbb{N}} \sum_{n=1}^{\infty} f_n d\mu = \sum_{n=1}^{\infty} (\int_{\mathbb{N}} f_n d\mu)$$

therefore

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} f_{nm} = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} f_{nm}$$

In analysis I, we used uniform convergence to get this result. At the moment we don't have negative f. Later we can retreat the same result.

Lemma 171. *(Fatou's) $f \geq 0$ measurable (not necessary convergent)*

$$\int_X \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int_X f_n d\mu$$

Proof.

$$\liminf f_n = \lim_{n\to\infty} \inf_{m\geq n} f_m$$

let $g_n = \inf_{m\geq n} f_m$, then $g_n \leq g_{n+1}$, g_n measurable.

Use monotone convergence

$$\lim_{n\to\infty} \int_X g_n d\mu = \int_X \lim_{n\to\infty} g_n d\mu$$

$$\int g_n d\mu \leq \int f_n d\mu \implies \lim_{n\to\infty} \int_X g_n d\mu \leq \liminf_{n\to\infty} \int_X f_n d\mu$$

□

Proposition 172. *$f \geq 0$ measurable*

$$\begin{aligned} \phi : \mathcal{M} &\to [0, \infty] \\ E &\mapsto \int_E f d\mu \end{aligned}$$

1) ϕ is a measure

2)

$$\int_E g d\phi = \int_E g f d\mu$$

if $g \geq 0$ measurable.

Proof. 1)

$$\begin{aligned} \phi(\coprod_{i=1}^{\infty} E_i) &= \int_{\coprod E_i} f d\mu = \int_X f \chi_{\coprod E_i} d\mu = \int_X f(\sum_{i=1}^{\infty} \chi_{E_i}) d\mu \\ &= \sum_{n=1}^{\infty} \int f \chi_{E_i} d\mu \because \text{monotone convergence} \\ &= \sum_{n=1}^{\infty} \int_{E_i} f d\mu = \sum_{n=1}^{\infty} \phi(E_i) \end{aligned}$$

2) standard method, use all the time. Prove for characteristic function, then use linearity to get for simple functions, then use monotone to get to all f.

Assume that $g = \chi_A$

$$\begin{aligned}\int_E g d\phi &= \phi(A \cap E) = \int_{A\cap E} f d\mu \\ &= \int_E f\chi_A d\mu = \int_E f g d\mu\end{aligned}$$

so it's true for χ_A, so true for positive measurable simple functions.

Fix $g \geq 0$ measurable, find $s_n \nearrow g$, use monotone convergence. $\square$

4.4 Dominated Convergence

Lecture 20 (4/9/13)

Definition 173. $L^1(\mu)$ Lebesgue integrable function. (X, m, μ) f a complex valued measurable function. $f \in L^1(\mu)$ iff

$$\int_X |f| d\mu < \infty$$

If $f = u + iv$

$$\int_X f d\mu := \int_X u d\mu + i \int_X v d\mu$$

$$u^{\pm} \leq |u| \leq |f| \qquad v^{\pm} \leq |v| \leq |f|$$

so

$$f \in L^1(\mu) \implies \int_X u^{\pm} d\mu < \infty \qquad \int_X u^{\pm} d\mu < \infty$$

The definition of Lebesgue integrable is somehow restricted.

Example 174. By Riemann

$$\int_{-\infty}^{\infty} \frac{\sin x}{x} dx = \lim_{M\to\infty} \int_{-M}^{M} \frac{\sin x}{x} dx = \pi$$

But in Lebesgue

$$\int_{-\infty}^{\infty} \left(\frac{\sin x}{x}\right)^+ dx = \infty$$

not integrable in Lebesgue.

Proposition 175. *$f, g \in L^1(\mu)$ and $\alpha, \beta \in \mathbb{C}$, then*

$$\alpha f + \beta g \in L^1(\mu) \text{ and } \int_X (\alpha f + \beta g) d\mu = \alpha \int_X f d\mu + \beta \int_X g d\mu$$

Proof. Assume f, g real (if complex, separate into u, v)

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

(note: it is not true $(f+g)^+ = f^+ + g^+$), then

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

use linearity of positive functions, so $f + g \in L^1(\mu)$.

$\alpha \in \mathbb{R}$, if $\alpha \geq 0$

$$\int \alpha f d\mu = \int (\alpha f)^+ - \int (\alpha f)^- = \alpha \int f^+ - \alpha \int f^-$$

if $\alpha \leq 0$

$$(\alpha f)^+ = (-\alpha) f^- \qquad (\alpha f)^- = (-\alpha) f^+$$

so

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

□

Theorem 176. *(Lebesgue dominated convergence) f_n a sequence of measurable functions, s.t.*

$$|f_n(x)| \leq g(x) \text{ and } g \in L^1(\mu)$$

assume that $f_n \to f$ pointwise, then $f \in L^1(\mu)$ and

$$\int_X |f_n - f| d\mu \to 0 \text{ and } \lim_{n \to \infty} \int f_n d\mu = \int_X f d\mu$$

This allows negative functions to come to play.

Proof. $|f| \leq g \implies \int_X |f| d\mu < \infty \implies f \in L^1(\mu)$.

$$|f_n - f| \leq 2g \implies 2g - |f_n - f| \geq 0$$

Use Fatou's

$$\int_X \liminf_{n\to\infty}(2g - |f_n - f|) \leq \liminf_{n\to\infty} \int_X (2g - |f_n - f|)d\mu$$

so

$$0 \leq \liminf(-\int_X |f_n - f|d\mu) \quad = \quad \lim_{n\to\infty} \inf_{n\geq m} -\int_X |f_n - f|d\mu = -\limsup \int_X |f_n - f|d\mu$$

Hence

$$\limsup \int_X |f_n - f|d\mu \leq 0$$

so $\lim_{n\to\infty} \int_X |f_n - f|d\mu = 0$ or $0 \leq \left|\int_X (f_n - f)d\mu\right| \leq \int_X |f_n - f|d\mu \to 0$, therefore

$$\int_X f_n d\mu \to \int_X f d\mu$$

□

5 Measure Theory

5.1 Measure Zero

Lecture 21 (4/16/13)

We have $f, g \in L^1(\mu)$ on X and $f(x) = g(x)\ \forall\, x \in E \subset X$ s.t.

$$\mu(E^c) = 0 \implies \int_X f d\mu = \int_X g d\mu$$

or

$$\int_X = \int_E + \int_{E^c} = \int_E$$

Definition 177. P is a proposition depending on $x \in X$, P is true almost everywhere wrt μ (denote *a.e.*-μ) iff

$$E = \{x \text{ s.t. } P(x) \text{ is not true}\} \implies \mu(E) = 0$$

In particular if $E = \{x \in X \text{ s.t. } f(x) \neq g(x)\}$, $\mu(E) = 0 \implies f = g$ *a.e.*-μ

Remark 178. If $f = g$ a.e.-$\mu \implies \int_X f d\mu = \int_X g d\mu$

It would be reasonable to ask if we have E s.t. $\mu(E) = 0$, for $F \subset E$ is F measurable? Is $\mu(F) = 0$?

if not, then $\int_E f d\mu = 0$ what happen to $\int_F f d\mu$?

What about $A \subset B$, A,B measurable and $\mu(A - B) = 0$, for E

$$A \subset E \subset B$$

Is E measurable?

ANS: Not true that E is measurable.

Definition 179. (X, m, μ) s.t. if $A, B \in m$, $A \subset B$ and $\mu(A - B) = 0$ then $\forall\, E$ s.t.

$$A \subset E \subset B \implies E \in m$$

we say that (X, m, μ) is complete.

It turns out that not complete space is very normal. Even we start with a complete space, after some operation the space becomes not complete (e.g. product space).

The only common way to complete the space is by extending measures to any subsets of measurable sets.

Example 180. Lebesgue measure $(\mathbb{R}, \mathcal{M}, m)$ complete but

$$(\mathbb{R} \times \mathbb{R}, ?, m \times m)$$

not complete.

Proposition 181. *(X, m, μ) $E \subset X$, $E \subset m^*$ iff $A \subset E \subset B$ and $A, B \in m$ and $\mu(A - B) = 0$ then define*

$$\mu(E) := \mu(A)$$

then (X, m^*, μ) is a complete measure space.

This is called completion of space.

Proof. 1) μ is well defined. If

$$A \subset E \subset B \text{ and } A' \subset E \subset B'$$

$A, B, A', B' \in m$ and $\mu(B - A) = \mu(B' - A') = 0$

$$\text{Is } \mu(A) = \mu(A')?$$

$$A - A' \subset E - A' \subset B' - A' \implies \mu(A - A') = 0$$
$$A' - A \subset E - A \subset B - A \implies \mu(A' - A) = 0$$

Then

$$\mu(A) = \mu(A \cap A') = \mu(A')$$

2) m^* is a σ algebra.

$$A \subset E \subset B \text{ and } \mu(B - A) = 0$$

so

$$B^c \subset E^c \subset A^c \text{ and } \mu(A^c - B^c) = 0$$

so

$$E^c \in m^*$$

since $\mu(B_n - A_n) = 0$

$$A_n \subset E_n \subset B_n \implies \cup A_n \subset \cup E_n \subset \cup B_n$$

then

$$\mu(\cup(B_n - \cup A_n)) \leq \mu(\cup(B_n - A_n)) \leq \sum_n \mu(B_n - A_n) = 0$$

so

$$\cup E_n \in m^*$$

Furthermore if $E_n \cap E_m = \varnothing$ $n \neq m$ then A_n are disjoint, that is why we define $\mu(E)$ to be the A's measure not B's.

$$\mu(\cup E_n) = \mu(\cup A_n) = \sum \mu(A_n) = \sum \mu(E_n)$$

□

How to calculate the measure of Lebesgue sets?

First use Fubini-Tonelli, slide volume into area $n \to n-1 \to n-2...$

then 1-D Riemann is Lebesgue, so we get the measure.

Proposition 182. *(X, m, μ) f_n measurable and f_n defined a.e. if*

$$\sum_{n\in\mathbb{N}} \int_X |f_n| d\mu < \infty \implies \sum_{n\in\mathbb{N}} f_n \text{ converges a.e.}$$

to a function $f(x)$; then

$$f(x) \in L^1(\mu) \quad \text{and } \textstyle\int_X f d\mu = \sum_{n\in\mathbb{N}} \int_X f_n d\mu$$

Remark 183. 1) use counting measure with these assumptions we can get swapping integral.

2) the same assumption we will use to prove Fubini-Tonelli.

Proof. f_n is defined on S_n, $\mu(S_n^c) = 0$, let

$$S = \bigcap_{n \in \mathbb{N}} S_n$$

every f_n is defined on S

$$\mu(S^c) = \mu\left((\cap S_n)^c\right) = \mu(\cup S_n^c) \leq \sum \mu(S_n^c) = 0$$

$$\sum_{n \in \mathbb{N}} \int_S |f_n| d\mu = \int_S \sum_{n \in \mathbb{N}} |f_n| d\mu < \infty$$

(Note $\int_X g d\mu < \infty \implies \mu(\{x \in X; g(x) = \pm\infty\}) = 0$)

$$E = \{x \in S; \sum |f_n(x)| < \infty\}$$

$$\mu(S - E) = 0 \implies \mu(E^c) = 0$$

$$g(x) = \sum_{n \in \mathbb{N}} |f_n(x)| \text{ for } x \in E$$

is measurable non negative defined a.e., so $g(x) \in L^1(\mu)$ and $\sum f_n(x)$ converges in E to a function $f(x)$, let

$$F_n(x) = \sum_{k=0}^{n} f_k$$

$F_n \to f$ a.e. and $|F_n| \leq \sum_{k=0}^{n} |f_k| \leq g$and $g \in L^1$, by dominated convergence $f \in L^1(\mu)$, so

$$\int_X f d\mu = \sum_{n=0}^{\infty} \int_X f_n d\mu$$

□

Corollary 184. *If μ counting measure and*

$$\sum_{i,j} |a_{ij}| < \infty \implies \sum_i \sum_j a_{ij} = \sum_j \sum_i a_{ij}$$

5.2 Construct Measure

We are not going to construct Lebesgue measure. The construction is very long and easy to get lost. We are not going to follow Rudin, because he takes the dual approach, a bit long and more than we need.

Rudin

$$f \to \int_X f d\mu$$

linear functional. Fix X gives a linear functional $I_X(f)$. From which find μ a measure s.t.

$$I_X(f) = \int_X f d\mu$$

by the Riesz representation theorem.

We take another approach that actually construct measure. We don't start from σ algebra with a measure m, because it has ∞ stuff, we start from an algebra.

Project for construct a measure.

Lecture 22 (4/18/13)

1) $\mathcal{F}_0$ algebra with a measure μ. Assume additive, only finite unions, not ∞ unions

$$(X, \mathcal{F}_0, \mu) \text{ and } \mu(X) < \infty$$

WLOG $\mu(X) = 1$ (one can always normalize)

2) $\mathcal{G}$ a collection of subsets and $G \in \mathcal{G}$ if $\exists\, A_n \in \mathcal{F}_0$ $A_1 \subset A_2 \subset \ldots$ s.t.

$$\bigcup_{n\in\mathbb{N}} A_n = G$$

$\mu(A_n)$ is increasing and $\mu(A_n) \leq 1 \implies$ convergent, so define

$$\mu(G) = \lim_{n\to\infty} \mu(A_n)$$

3) Extend μ to $P(X)$, for $A \subset X$

$$\mu^*(A) = \inf\{\mu(G) : G \supset A \text{ and } G \in \mathcal{G}\}$$

where μ^* is not a measure.

What we get is that μ^* is an outer measure and we will show the following soon

i) $\mu^*(\varnothing) = 0$

ii) $A \subset B \implies \mu^*(A) \leq \mu^*(B)$

iii) A_n are disjoint $\mu^*(\cup_{n\in\mathbb{N}} A_n) \leq \sum_{n\in\mathbb{N}} \mu^*(A_n)$ (called countable subadditivity)

iv) $\mathcal{H} = \{H \in P(X); \mu^*(H) + \mu^*(H^c) = 1\}$

v) X is σ finite.

$(X, \mathcal{H}, \mu^*)$ is a measure space, μ^* is a measure and $\mathcal{F}_0 \subset \mathcal{G} \subset \sigma(\mathcal{F}_0) \subset H$. H is the completion of $\sigma(\mathcal{F}_0)$.

Example 185. Prototype: $\mathcal{F}_0$ are finite union of n-cells, $\mathcal{G}$ open sets, $\sigma(\mathcal{F}_0)$ Borel, $\mathcal{H}$ completion of Borel.

Lemma 186. *$\mathcal{F}_0$ algebra.*

$$A_n \nearrow A,\ A_n \in \mathcal{F}_0 \textit{ and } A \subset X,$$

$$A'_n \nearrow A',\ A'_n \in \mathcal{F}_0 \textit{ and } A' \subset X,$$

If $A \subset A'$ then

$$\lim_{n\to\infty} \mu(A_n) \leq \lim_{n\to\infty} \mu(A'_n)$$

In particular if $A = A' \implies \lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \mu(A'_n)$.

Proof. Fix n, m, so $A_n \cap A'_m \subset A_n \cap A'_{m+1}$, $A_n \subset A \subset A'$, and $A_n \cap A'_m \nearrow A_n \cap A' = A_n$

(in the proof we don't need countable additive)

$$\lim_{m\to\infty} \mu(A_n \cap A'_m) = \mu(\bigcup_{m\in\mathbb{N}} (A_n \cap A'_m)) = \mu(A_n)$$

Do $\lim_{m\to\infty} (\mu(A'_m) \geq \mu(A_n \cap A'_m))$, then do $\lim_{n\to\infty} (\lim_{m\to\infty} \mu(A'_m) \geq \mu(A_n))$, so

$$\lim_{m\to\infty} \mu(A'_m) \geq \lim_{n\to\infty} \mu(A_n)$$

□

Proposition 187. *Given $\mathcal{F}_0$, make $\mathcal{G}$ the set of increasing sequences in $\mathcal{F}_0$. Define*

$$\mu(G) := \lim_{n\to\infty} \mu(A_n)$$

if $A_n \nearrow G$, $A_n \in \mathcal{F}_0$

1) $\mathcal{F}_0 \in \mathcal{G}$ and $\mu(G) \leq 1$

2) $G_1 \subset G_2 \implies \mu(G_1) \leq \mu(G_2)$

3) $\mu(G_1) + \mu(G_2) = \mu(G_1 \cup G_2) + \mu(G_1 \cap G_2)$

4) $G_n \subset G_{n+1} \implies G = \bigcup_{n=1}^{\infty} G_n \in \mathcal{G}$ and $\mu(G_n) \to \mu(G)$

Proof. 1) obvious

2) By the lemma, $A_n \nearrow G_1$,

$$A'_n \nearrow G_2 \implies \lim_{n\to\infty} \mu(A_n) \leq \lim_{n\to\infty} \mu(A'_n)$$

if $G_1 \subset G_2$.

3) $A_n \nearrow G_1$, $A'_n \nearrow G_2 \implies A_n \cup A'_n \nearrow G_1 \cup G_2$ and $A_n \cap A'_n \nearrow G_1 \cap G_2$

$$\lim_{n\to\infty} \big(\mu(A_n) + \mu(A'_n) = \mu(A_n \cup A'_n) + \mu(A_n \cap A'_n)\big)$$

Hence

$$\mu(G_1) + \mu(G_2) = \mu(G_1 \cup G_2) + \mu(G_1 \cap G_2)$$

4) Similar to the proof of one of the preliminary lemma Arzela-Ascoli. Since

$$\begin{array}{ccc} A_{11} \subset A_{12} \subset A_{13}... & & G_1 \\ & & \cap \\ A_{21} \subset A_{22} \subset A_{23}... & & G_2 \\ \vdots & & \vdots \end{array}$$

Let $D_1 = A_{11}$, $D_2 = A_{12} \cup A_{22}$, $D_3 = A_{13} \cup A_{23} \cup A_{33}$, etc

$$D_n = \bigcup_{i=1}^{n} A_{in}$$

$$A_{nm} \subset D_m \subset G_m$$

if $n \leq m$. Take $\cup_{m\in\mathbb{N}}$ above

$$\underbrace{\bigcup_{m=1}^{\infty} A_{nm}}_{G_n} \subset \bigcup_m D_m \subset \bigcup_m G_m = G$$

Take $\cup_{n\in\mathbb{N}}$ above, get

$$G \subset \bigcup_m D_m \subset G \implies \bigcup_{m=1}^{\infty} D_m = G$$

$$\mu(A_{nm}) \leq \mu(D_m) \leq \mu(G_m)$$

take $\lim_{m\to\infty}$ above, then take $\lim_{n\to\infty}$, get

$$\lim_{n\to\infty} \Big(\mu(G_n) \leq \mu(G) \leq \lim_{m\to\infty} \mu(G_m)\Big)$$

so

$$\mu(G) = \lim_{m\to\infty} \mu(G_m)$$

$\square$

Proposition 188. *$\mathcal{G}$, μ are in the previous proposition. Define μ^* on $\mathcal{P}(X)$,*

$$\mu^*(A) = \inf\{\mu(G);\ A \subset G \text{ and } G \in \mathcal{G}\}$$

1) $\mu^*(\varnothing) = 0$, $\mu^*(A) \le 1$,$\mu^*(G) = \mu(G)$ if $G \in \mathcal{G}$

2) $\mu^*(A) + \mu^*(B) \ge \mu^*(A + B) + \mu^*(A \cap B)$

3) $A \subset B \implies \mu^*(A) \le \mu^*(B)$

4) $A_n \nearrow A \implies \mu^*(A_n) \to \mu^*(A)$

5) $\mu^*(\coprod_{n\in\mathbb{N}} A_n) \le \sum_{n=1}^{\infty} \mu^*(A_n)$

This shows μ^* indeed is an outer measure. 5) indicates it is not a measure.

Proof. 1) obvious

2) Fix $\epsilon > 0$ find $G_1 \supset A$ and $G_2 \supset B$, G_1, $G_2 \in \mathcal{G}$ s.t.

$$\mu^*(A) + \epsilon \ge \mu^*(G_1) = \mu(G_1),\ \mu^*(B) + \epsilon \ge \mu^*(G_2) = \mu(G_2)$$

$$\begin{aligned} \mu^*(A) + \mu^*(B) + 2\epsilon &\ge \mu(G_1) + \mu(G_2) \\ &= \mu(G_1 \cup G_2) + \mu(G_1 \cap G_2) \\ &\ge \mu^*(A \cup B) + \mu^*(A \cap B) \end{aligned}$$

(so we get

$$\mu^*(A) + \mu^*(A^c) \ge 1 \tag{5.1}$$

for free)

3) because $G \supset B \implies G \supset A$

4) $A_n \subset A_{n+1}$ and $A = \cup_{n\in\mathbb{N}} A_n \supset A_n$

$$\lim \mu^*(A_n) \le \mu^*(A)$$

Fix $\epsilon > 0$ find $G_n \in \mathcal{G}$ s.t. $A_n \subset G_n$,

$$\mu^*(A_n) + \epsilon 2^{-n} \ge \mu^*(G_n) \text{ and } A \subset \bigcup_{n\in\mathbb{N}} G_n$$

$$\mu^*(A) \leq \mu^*(\bigcup_{n\in\mathbb{N}} G_n) = \mu^*(\lim_{n\to\infty}\bigcup_{k=1}^{n} G_k) = \lim_{n\to\infty}\mu^*(\bigcup_{k=1}^{n} G_k)$$

It is enough to show

$$\mu(\bigcup_{k=1}^{n} G_k) \leq \mu^*(A_n) + \epsilon\sum_{k=1}^{n} 2^{-k}$$

indeed if $n = 1$, it's true, induction

$$\begin{aligned}
\mu(\bigcup_{k=1}^{n+1} G_k) &= \mu(\bigcup_{k=1}^{n} G_k) + \mu(G_{n+1}) - \mu(\bigcup_{k=1}^{n} G_k \cap G_{n+1}) \\
&\leq \mu^*(A_n) + \epsilon\sum_{k=1}^{n} 2^{-k} + \mu^*(A_{n+1}) + \epsilon 2^{-(n+1)} - \mu(\bigcup_{k=1}^{n} G_k \cap G_{n+1})
\end{aligned}$$

$$\mu^*(\bigcup_{k=1}^{n} G_k \cap G_{n+1}) \geq \mu^*(G_n \cap G_{n+1}) \geq \mu^*(A_n \cap A_{n+1}) = \mu^* A_n,$$

implies

$$\mu(\bigcup_{k=1}^{n+1} G_k) \leq \mu^*(A_{n+1}) + \epsilon\sum_{k=1}^{n+1} 2^{-k}$$

5) By 2)

$$\mu^*(\lim_{n\to\infty}\coprod_{k=1}^{n} A_k) = \lim_{n\to\infty}\mu^*(\coprod_{k=1}^{n} A_k) \leq \lim_{n\to\infty}\sum_{k=1}^{n}\mu^*(A_k)$$

□

Lecture 23 (4/23/13)

We have done so far: $\mathcal{F}_0$ algebra, $\mathcal{G} \supset \mathcal{F}_0$ the set of countable union of elements in $\mathcal{F}_0$, μ a finite measure on $\mathcal{F}_0$, also a finite measure on $G \in \mathcal{G}$, by setting $A_n \nearrow G$, $A_n \in \mathcal{F}_0$, $\mu(G) = \lim_{n\to\infty}\mu(A_n)$, and for $A \subset X$, we let

$$\mu^*(A) = \inf\{\mu(G); G \supset A, G \in \mathcal{G}\}$$

which is an outer measure.

Proposition 189. *$(X, \mathcal{F}_0, \mu)$ and μ^* is the outer measure associated to μ,*

$$\mathcal{H} = \{H \subset X; \mu^*(H) + \mu^*(H^c) = 1\}$$

1) $\mathcal{G} \subset \mathcal{H}$

2) $\mathcal{H}$ is a σ algebra

3) μ^* is a measure on $\mathcal{H}$ that extends μ. $(X, \mathcal{H}, \mu^*)$ is a measurable space.

This shows although μ^* is not a measure on $\mathcal{P}(X)$, it is a measure on $\mathcal{H}$, and by 2) μ^* is a measure on $\sigma(\mathcal{F}_0)$, too. Application of this allows us to work with $\mathbb{R}^n$, where no finite measure exists.

Proof. 1) want to show $G \in \mathcal{G} \implies \mu^*(G) + \mu^*(G^c) = 1$.

In general $G^c \notin \mathcal{G}$. Let$A_n \nearrow G$; $A_n \in \mathcal{F}_0$

$$A_n \subset G \implies G^c \subset A_n^c \implies \mu^*(G^c) \leq \mu^*(A_n^c)$$

$\lim_{n\to\infty} (\mu(A_n) + \mu^*(G^c) \leq \mu(A_n) + \mu^*(A_n^c) = 1)\ \forall n$, so

$$\mu^*(G) + \mu^*(G^c) \leq 1$$

On the another hand

$$\mu^*(A \cup B) + \mu^*(A \cap B) \leq \mu^*(A) + \mu^*(B)$$

set $A = G$, $B = G^c$

$$1 + 0 \leq \mu^*(G) + \mu^*(G^c)$$

2) Prove that $\mathcal{H}$ is an algebra.

$$H \in \mathcal{H} \implies H^c \in \mathcal{H}.$$

H_1,$H_2 \in \mathcal{H}$,

$$\mu^*(H_1 \cup H_2) + \mu^*(H_1 \cap H_2) \leq \mu^*(H_1) + \mu^*(H_2) \tag{5.2}$$

$$\mu^*(H_1^c \cup H_2^c) + \mu^*(H_1^c \cap H_2^c) \leq \mu^*(H_1^c) + \mu^*(H_2^c) \tag{5.3}$$

combining (5.2), (5.3) above, by (5.1)

$$[\underbrace{\mu^*(H_1 \cup H_2) + \mu^*((H_1 \cup H_2)^c)}_{\geq 1}] + [\underbrace{\mu^*(H_1 \cap H_2) + \mu^*((H_1 \cap H_2)^c)}_{\geq 1}] \leq 2$$

so

$$\mu^*(H_1 \cup H_2) + \mu^*((H_1 \cup H_2)^c) = 1 \implies H_1 \cup H_2 \in \mathcal{H}$$

same thing for $H_1 \cap H_2$. So $\mathcal{H}$ is an algebra.

Prove that H contains countable unions

$$H_n \in \mathcal{H}$$

we can assume that $H_n \subset H_{n+1}$

$$\mu^*(H_n) \to \mu^*(H)$$

$H = \bigcup_{n\in\mathbb{N}} H_n$, (when we proved subadditivity, we actually got a little bit more.)

$$H_n \subset H \implies H^c \subset H_n^c \implies \mu^*(H^c) \leq \mu^*(H_n^c)$$

$$\lim_{n\to\infty} (\mu^*(H_n) + \mu^*(H^c) \leq \mu^*(H_n) + \mu^*(H_n^c) = 1)$$

$$\mu^*(H) + \mu^*(H^c) \leq 1 \implies H \in \mathcal{H}$$

3) Prove that μ^* is finitely additive. Suppose not, i.e. $\exists\, A$, B s.t.

$$\mu^*(A \cup B) < \mu^*(A) + \mu^*(B)$$

then (5.2), (5.3) give

$$2 \leq [\mu^*(H_1 \cup H_2) + \mu^*((H_1 \cup H_2)^c)] + [\mu^*(H_1 \cap H_2) + \mu^*((H_1 \cap H_2)^c)] < 2$$

contradiction.

Since $\mathcal{H}$ is σ algebra, μ^* is σ additive on $\mathcal{H}$, that is for $H_n \in \mathcal{H}$ disjoint

$$\mu^*(\cup_{n\in\mathbb{N}} H_n) = \mu^*(\lim_{n\to\infty} \cup_{k=1}^n H_k) = \lim_{n\to\infty} \mu^*(\cup_{k=1}^n H_k) = \sum_{n=1}^{\infty} \mu^*(H_n)$$

so μ^* is a measure. $\square$

Turn out $\mathcal{H}$ is the completion of G.

Proposition 190. $(X, \mathcal{H}, \mu)$ $\mathcal{H}$ *Lebesgue measure over* $\mathcal{F}_0$

1) $\mathcal{H}$ *is complete*

2) $\mathcal{H}$ *is the completion of* $\sigma(\mathcal{F}_0)$ *denote*

$$\mathcal{H} \equiv \widetilde{\sigma(\mathcal{F}_0)}$$

i.e. the smallest σ *algebra contains* σ *algebra of* $\mathcal{F}_0$*.*

e.g. in $\mathbb{R}^n$, $\mathcal{H} =$ completion of Borel spaces.

Proof. 1) $H \in \mathcal{H}$ s.t. $\mu(H) = 0$. then

$$E \subset H \implies E \in \mathcal{H}$$

Indeed

$$\mu^*(E) \leq \mu^*(H) = \mu(H) = 0 \implies \mu^*(E) = 0$$
$$H^c \subset E^c \implies 1 = \mu(H^c) \leq \mu^*(E^c) \leq 1$$
$$\mu^*(E) + \mu^*(E^c) = 0 + 1 = 1 \implies E \in \mathcal{H}$$

2) $\widetilde{\sigma(\mathcal{F}_0)} \subset \mathcal{H}$ because $\mathcal{H}$ is complete and $\widetilde{\sigma(\mathcal{F}_0)}$ is denoted as the smallest completion.

Want to show $\mathcal{H} \subset \widetilde{\sigma(\mathcal{F}_0)}$, i.e. $A, B \in \sigma(\mathcal{F}_0)$ and $A \subset B$, $A \subset H \subset B$ and $\mu(A - B) = 0$ show $H \in \widetilde{\sigma(\mathcal{F}_0)}$.

Fix $H \in \mathcal{H}$, find $G_n \in \mathcal{G} \subset \sigma(\mathcal{F}_0)$ s.t. $H \subset G_n$ and $\mu(G_n - H) < \frac{1}{n}$, assume that

$$G_{n+1} \subset G_n$$

Let $G = \bigcap_{n=1}^{\infty} G_n \implies H \subset G$, and

$$\mu(G - H) = \lim_{n \to \infty} \mu(G_n - H) = 0$$

Find $H_n \in \mathcal{G}$, $H_n^c \in \sigma(\mathcal{F}_0)$, $H^c \subset H_n$, and $\mu(H_n - H^c) < \frac{1}{n}$

$$H_{n+1} \subset H_n$$

Let$G' = \cup_{n \in \mathbb{N}} H_n^c$, $\quad G' \in \sigma(\mathcal{F}_0)$, $G' \subset H$, and

$$\mu(H - G') = \lim_{n \to \infty} (H - H_n^c) = 0$$

continuity from below. Hence

$$G' \subset H \subset G$$

and $\mu(G - H) = \mu(H - G') = \mu(G - G') = 0$ $\square$

5.3 Caratheodory Extension

Now we are going to extend the result to measure that is not finite.

Definition 191. $(X, \mathcal{F}_0, \mu)$ ($\mathcal{F}_0$ an algebra or a σ algebra) μ is σ finite measure wrt $\mathcal{F}_0$ if $\exists\, A_n \in \mathcal{F}_0$ s.t.

$$\bigcup_{n=1}^{\infty} A_n = X \text{ and } \mu(A_n) < \infty$$

Remark 192. Whenever countable union is encountered, we can always rewrite it as union of disjoint set, so assume A_n disjoint in the definition. Typically to define a measure start from $\mathcal{F}_0$ an algebra then get σ algebra $\sigma(\mathcal{F}_0)$.

It can happen that μ is σ finite wrt $\sigma(\mathcal{F}_0)$ and not σ finite wrt $\mathcal{F}_0$. The subtle difference is the key to the extension. From a set theoretical perspective, $\sigma(\mathcal{F}_0)$ contains more sets so easier to cover by finite sets.

Example 193. what can go wrong? $\mathbb{Q}$. Make $\mathcal{F}_0$ using $(a, b]$, (a, ∞), $(-\infty, b]$, $(-\infty, \infty)$, $\varnothing \subset \mathbb{Q}$, where intervals means only points in $\mathbb{Q}$. $\sigma(\mathcal{F}_0)$ contains single point

$$\bigcap_{n \in \mathbb{N}} (a_n, b] = \{b\}$$

$a_n \nearrow b$, so

$$\sigma(\mathcal{F}_0) = \mathcal{P}(\mathbb{Q})$$

because $\mathbb{Q}$ is countable.

Suppose μ is the counting measure on $\mathcal{P}(\mathbb{Q})$, for $A \in \mathcal{F}_0$

$$\mu(A) = \begin{cases} \infty & A \neq \varnothing \implies \text{not}\, \sigma \text{ finite on } \mathcal{F}_0 \\ 0 & A = \varnothing \end{cases}$$

However $\{p\} \in \sigma(\mathcal{F}_0) \implies \mu(\{p\}) = 1 \implies \mu$ is σ finite on $\sigma(\mathcal{F}_0)$.

Consequently it has no unique extension. If $\tilde{\mu}_1$ is the counting measure

$$\tilde{\mu}_1|_{\mathcal{F}_0} = \mu$$

take $\tilde{\mu}_2 = 2\tilde{\mu}_1$ also an extension.

This sums up in Caratheodory

Lecture 24 (4/25/13)

Theorem 194. *(Caratheodory extension)* $(X, \mathcal{F}_0, \mu)$ *$\mathcal{F}_0$ an algebra, μ a measure. μ is σ finite wrt $\mathcal{F}_0$, then $\exists\, \lambda$ a measure, extension of μ, on $\sigma(\mathcal{F}_0)$, i.e.*

$$\lambda|_{\mathcal{F}_0} = \mu$$

We already know if μ is finite on $\mathcal{F}_0$, it has extension. First half of Caratheodory says μ, σ finite, has extension. The second half says the extension is unique.

Proof. $X = \bigcup_{n\in\mathbb{N}} A_n$ s.t. $A_n \in \mathcal{F}_0$ and $\mu(A_n) < \infty$ we can assume that A_n are disjoint if not replace A_n with

$$(A_n - A_1) - A_2 - \ldots - A_{n-1}$$

For $B \in \mathcal{F}_0$, let

$$\mu_n(B) = \mu(B \cap A_n)$$

so

$$\mu_n(B) \leq \mu(A_n) < \infty$$

use the extension for finite measure μ_n, and define λ_n on $\sigma(\mathcal{F}_0)$ s.t.

$$\lambda_n|_{\mathcal{F}_0} = \mu_n$$

Then

$$\lambda = \sum_n \lambda_n$$

is the desired extension. Indeed if $B \in \mathcal{F}_0$, since A_n are disjoint, $B = \bigcup_{n\in\mathbb{N}} B \cap A_n$. Hence

$$\begin{aligned}
\lambda(B) &= \sum_n \lambda_n(B) = \sum_n \mu_n(B) \\
&= \sum_n \mu(B \cap A_n) = \mu(B \cap (\bigcup A_n)) = \mu(B)
\end{aligned}$$

If $E_n \in \sigma(\mathcal{F}_0)$ disjoint,

$$\begin{aligned}
\lambda(\bigcup_{n\in\mathbb{N}} E_n) &= \sum_{n\in\mathbb{N}} \lambda_n(\bigcup_{k\in\mathbb{N}} E_k) \\
&= \sum_{n\in\mathbb{N}} \sum_{k\in\mathbb{N}} \lambda_n(E_k) \because \text{they are all positive, swap allowed} \\
&= \sum_{k\in\mathbb{N}} \sum_{n\in\mathbb{N}} \lambda_n(E_k) = \sum_{k\in\mathbb{N}} \lambda(E_k)
\end{aligned}$$

□

Definition 195. $C \subset \mathcal{P}(X)$ is a monotone class iff

1) $\forall\, A_n \in C,\ A_n \subset A_{n+1} \implies \bigcup_{n\in\mathbb{N}} A_n \in C$

2) $\forall\, A_n \in C$ s.t. $A_n \supset A_{n+1} \implies \bigcap_{n\in\mathbb{N}} A_n \in C$

Remark 196. $\mathcal{F}_0$ a algebra $\exists$ a smallest monotone class containing it

1) show that the intersection of monotone class is a monotone class

2) $\sigma(\sigma(\mathcal{F}_0)) \supset \sigma(\mathcal{F}_0) \supset \mathcal{F}_0$ is monotone

Theorem 197. *(monotone class) $\mathcal{F}_0$ is an algebra, the smallest monotone class containing $\mathcal{F}_0$ is $\sigma(\mathcal{F}_0)$.*

Usually it is easier to prove something is monotone class than to prove it is σ algebra.

Proof. Let M be the smallest monotone class containing $\mathcal{F}_0$.$\mathcal{F}_0 \subset M \subset \sigma(\mathcal{F}_0)$. For $B \in M$, put

$$M_B = \{A \in M; A \cap B, A \cap B^c, A^c \cap B \in M\}$$

Show M_B is monotone.

Take $A_n \in M$ s.t. $A_n \cap B, A_n \cap B^c, A_n^c \cap B \in M$ and $A_n \subset A_{n+1}$, need to show $A = \cup A_n \in M$

claim $A \cap B \in M$, indeed

$$(\cup A_n) \cap B = \cup(A_n \cap B)$$

so

$$A_n \cap B \subset A_{n+1} \cap B \implies \cup(A_n \cap B) \in M$$

Now say the sequence is decreasing, $A_n \supset A_{n+1} \qquad A = \bigcap_{n\in\mathbb{N}} A_n$, then

$$A \cap B = \cap A_n \cap B \implies A \cap B \in M$$

Similar to check other 2 properties.

Show $M = M_B$.

Suppose that $B \in \mathcal{F}_0 \implies M_B \supset \mathcal{F}_0$. If $B \notin \mathcal{F}_0$,$\forall\, A \in \mathcal{F}_0$, $A \cap B, A \cap B^c, A^c \cap B \in M$ still imply $M_B \supset \mathcal{F}_0$. Because M is the smallest monotone class containing $\mathcal{F}_0$, $M_B \supset M$. But by the definition of M_B, $M_B \subset M \implies M_B = M$.

M is an algebra.

$$\forall A, B \in M,\ A \cap B, A \cap B^c, A^c \cap B \in M$$

shows it has finite unions. Since M contains increasing sequences, and use $A \cap B, A \cap B^c, A^c \cap B$, it contains countable unions, so it's σ algebra.

$$\mathcal{F}_0 \subset M \subset \sigma(\mathcal{F}_0)$$

because M smallest monotone class.

$$\mathcal{F}_0 \subset \sigma(\mathcal{F}_0) \subset M$$

because $\sigma(\mathcal{F}_0)$ is the smallest σ algebra. □

Theorem 198. *(Caratheodory uniqueness) $(X, \mathcal{F}_0, \mu)$ $\mathcal{F}_0$ an algebra, μ a measure. μ is σ finite wrt $\mathcal{F}_0$, then $\exists!\lambda$ a measure, extension of μ, on $\sigma(\mathcal{F}_0)$, i.e.*

$$\lambda|_{\mathcal{F}_0} = \mu \text{ and } \lambda = \mu^*$$

The proof is very standard, used in may other proofs.

Proof. $X = \bigcup_{n\in\mathbb{N}} A_n$, A_n disjoint and $A_n \in \mathcal{F}_0$, let

$$\lambda_n(B) = \lambda(B \cap A_n)$$

Fix n, if $B \in \mathcal{F}_0$, $\lambda_n(B) = \mu_n(B) = \mu_n^*(B)$. Let

$$\mathcal{C} = \{B \in \sigma(\mathcal{F}_0) \text{ s.t. } \lambda_n(B) = \mu_n^*(B)\}$$

so $\mathcal{F}_0 \subset \mathcal{C}$. Prove that $\mathcal{C}$ is monotone, take $B_k \in \mathcal{C}$ and $B_k \subset B_{k+1}$ $B_k \nearrow B = \bigcup B_k$, so

$$\lambda_n(B_k) = \mu^*(B_k) \implies \lambda_n(B) = \lim_{k\to\infty} \lambda_n(B_k)$$

and $\mu_n^*(B) = \lim_{k\to\infty} \mu_n^*(B_k)$, then

$$\lambda_n(B) = \mu_n^*(B) \implies B \in \mathcal{C}$$

take $B_k \supset B_{k+1}$ $\quad B_k \searrow B = \bigcap B_k$, use continuity from above

$$\lambda_n(B) = \mu_n^*(B) \implies B \in \mathcal{C}$$

so $\mathcal{C}$ is monotone.

$$C \supset \sigma(\mathcal{F}_0) \implies C = \sigma(\mathcal{F}_0)$$

That says
$$\forall\, n \in \mathbb{N}\ \lambda_n(B) = \mu_n^*(B)\ \forall\, B \in \sigma(\mathcal{F}_0)$$
hence
$$\sum \lambda_n = \sum \mu_n^* \implies \lambda = \mu^*$$

$\square$

Proposition 199. *$(X, \sigma(\mathcal{F}_0), \mu)$ μ is σ finite wrt $\mathcal{F}_0$. $\forall\, A \in \sigma(\mathcal{F}_0)$ s.t. $\mu(A) < \infty$ and $\forall\, \epsilon > 0\ \exists\, B \in \mathcal{F}_0$ s.t.*
$$\mu(A\Delta B) < \epsilon.$$

$A\Delta B$ symmetric difference, $(A - B) \cup (B - A)$.

Proof. Assume that μ is finite. $\mathcal{G}$ is the set of all countable unions of increasing sets in $\mathcal{F}_0$. $A \in \sigma(\mathcal{F}_0)$, fix $\epsilon > 0$ find $G \in \mathcal{G}$ s.t.
$$\mu(A) + \epsilon > \mu(G) \text{ and } G \supset A$$
$\exists\, B_n \in \mathcal{F}_0$ s.t. $B_n \nearrow G$ and $\mu(B_n) \to \mu(G)$. $\exists\, N_\epsilon$ s.t.
$$\mu(G) - \mu(B_{N_\epsilon}) < \epsilon$$
then
$$\begin{aligned} \mu(A\Delta B_{N_\epsilon}) &= \mu(A - B_{N_\epsilon}) + \mu(B_{N_\epsilon} - A) \\ &< \mu(G - B_{N_\epsilon}) + \mu(G - A) < 2\epsilon \end{aligned}$$
so we are done for the finite case.

The general case, we do it in 2 steps. $X = \bigcup A_n$ $A_n \in \mathcal{F}_0$ disjoint $\mu(A_n) < \infty$. Let
$$\mu_n(\cdot) = \mu(\cdot \cap A_n)$$
so it is a finite measure. Give $A \in \sigma(\mathcal{F}_0)$, find $B_n \in \mathcal{F}_0$ s.t. $\mu_n(A\Delta B_n) < \epsilon 2^{-n}$, that is
$$\begin{aligned} \mu_n(A\Delta B_n) &= \mu((A\Delta B_n) \cap A_n) \\ &= \mu((A\Delta B_n \cap A_n) \cap A_n) = \mu_n(A\Delta(B_n \cap A_n)) < \epsilon 2^{-n} \end{aligned}$$
Replace B_n with $B_n \cap A_n \in \mathcal{F}_0$. Assume that $B_n \subset A_n$, so B_n disjoint, than for $B = \bigcup_{n\in\mathbb{N}} B_n$
$$\mu(A\Delta B) = \sum_{n\in\mathbb{N}} \mu_n(A\Delta B) = \sum_{n\in\mathbb{N}} \mu_n(A\Delta B_n) < \epsilon$$

The 2nd equality is because

$$\begin{aligned}\mu_n(A\Delta B) &= \mu(A\Delta B\cap A_n) = \mu(A\Delta(B\cap A_n)\cap A_n)\\ &= \mu(A\Delta B_n\cap A_n) = \mu_n(A\Delta B_n)\end{aligned}$$

Note B is not necessary in $\mathcal{F}_0$, but $C_N = \cup_{n=1}^N B_n \in \mathcal{F}_0$

$$\mu(A\Delta C_N) = \mu(A - C_N) + \mu(C_N - A)$$

Clearly

$$A - \cup_{n=1}^N B_n \searrow A - B \text{ and } \cup_{n=1}^N B_n - A \nearrow B - A$$

use continuity, pick N large enough. Since $\mu(A) < \infty$,

$$\mu(A\Delta C_N) \leq \mu(A\Delta B) < \epsilon$$

□

5.4 Lebesgue-Stieltjes Measure

Lecture 25 (4/30/13)

Example 200. $\mathbb{R}$, let $\mathcal{F}_0$ be the algebra generated by $(a, b]$. F is a real increasing function, right continuous (RC). One can define a measure here like Stieltjes

$$\mu((a, b]) = F(b) - F(a)$$

Can extend it to $\sigma(\mathcal{F}_0)$? What is the outer measure $\mu((a, b)) =$?

$b_n \nearrow b$,$(a, b_n] \nearrow \cup(a, b_n] = (a, b)$, $F(b_n) \to F(b-)$,

summary:

$$\mu(a, b) = F(b-) - F(a)$$
$$\mu([a, b]) = F(b) - F(a-)$$
$$\mu([a, b)) = F(b-) - F(a-)$$
$$\mu(\{a\}) = F(a) - F(a-)$$

gives size of jump.

This type of measure is not the most general thing to do. but if we require the measure to be finite, this is the most general.

Definition 201. We call μ a Lebesgue-Stieltjes (LS) measure on $B(\mathbb{R})$ iff

$$\mu((a, b]) < \infty$$

Remark 202. $\forall\, \mu$ on $B(\mathbb{R})$ which is LS $\exists\, F$ real function increasing and RC s.t.

$$\mu((a,b]) = F(b) - F(a)$$

Proof. Fix the value of $F(0)$, define $F(x)$ via

$$\mu((0,x]) := F(x) - F(0)$$

It is absolute increasing, for

$$\begin{aligned} \mu((a,b]) &= \mu((0,b] - (0,a]) = F(b) - F(0) - (F(a) - F(0)) \\ &= F(b) - F(a) \geq 0 \end{aligned}$$

Pick $x_n \searrow x$, $(0,x_n] \searrow (0,x]$, so

$$\mu((0,x_n]) \to \mu((0,x])$$

i.e.$F(x_n) - F(0) \to F(x) - F(0) \implies F(x_n) \to F(x) \implies \lim_{t \to x^+} F(t) = F(x)$. $\square$

What about in higher dimension?

Definition 203. $a, b \in \mathbb{R}^n$

$$(a,b] = \{x \in \mathbb{R}^n; u_i \leq x_i \leq b_i \,\forall\, i = 1, ..., n\}$$

$\mathcal{F}_0$ is the algebra generated by $(a,b]$

$$\sigma(\mathcal{F}_0) = B(\mathbb{R}^n)$$

because we can write (a,b) as countable union of $(a,b_n]$ and (a,b) is a countable base of the topology.

Definition 204. μ on $B(\mathbb{R}^n)$ is LS iff $\mu((a,b]) < \infty$.

One can take $F(x)$ increasing right continuous

$$\mu((a,b)) = \prod_{i=1}^{n} (F(b_i) - F(a_i))$$

if $F = I$, $\mu((a,b))$ is the volume of the call. The general case is built with a function $F(x_1, ..., x_n)$. so we see the construction of the measure always take some limit.

Let us now do another approximation of measure, and we require the approximated sets as explicitly as we can get.

Proposition 205. $\mathbb{R}^k$ $B(\mathbb{R}^k)$ μ *is* σ *finite.*

1) $B \in B(\mathbb{R}^k)$

$$\mu(B) = \sup\{\mu(K); K \subset B \text{ and } K \text{is compact}\}$$

2) Assume also that μ *is LS, then*

$$\mu(B) = \inf\{\mu(U); U \supset B \text{ and } U \text{is open}\}$$

Remark 206. Addition assumptions in 2) is necessary, because there are examples of σ finite but of not LS.

Example 207. Say $\mathbb{R}$, Borel set. Algebra contains $(a, b]$,

$$\mu((b, \infty)) = F(\infty) - F(b) \implies F(\infty) = \infty$$

More interesting example below

Example 208.

$$S = \{\frac{1}{n}; n \in \mathbb{N}\} \subset \mathbb{R} \text{ with } \mu\{\frac{1}{n}\} = \frac{1}{n}$$

For $A \subset \mathbb{R}$, let $\mu(A) = \mu(A \cap S)$, show this is not Lebesgue

$$\mu([0, 1]) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

on the other hand, since

$$\mu(S^c) = 0 \text{ and } \mathbb{R} = \bigcup_{n \in \mathbb{N}} \{\frac{1}{n}\} \cup S^c$$

countable $\implies \sigma$ finite.

So if apply the proposition brutally,

$$\mu(\{0\}) = 0$$

but for any $0 \in U$ open

$$\mu(V) = \infty$$

Indeed $\exists\, \epsilon > 0$ s.t.$(-\epsilon, \epsilon) \subset U$, then $\mu(U) \geq \mu((-\epsilon, \epsilon)) = \sum_{k=n}^{\infty} \frac{1}{k} = \infty$, so $\inf\{\mu(U); U \supset \{0\}\} = \infty$.

Proof. 1) Assume that μ is finite. Ω is the collection of sets $B \in B(\mathbb{R}^n)$ that have the property in 1).

Use monotone class. Let $\mathcal{F}_0$ be the algebra generated by $(a, b]$, so $\sigma(\mathcal{F}_0) = B(\mathbb{R}^k)$. We know for $B \in \mathcal{F}_0$, we can find compact cell $[a_n, b] \nearrow (a, b]$ and $[a_n, b] \subset (a, b]$. Continuous from below implies

$$\mu([a_n, b]) \to \mu((a, b])$$

so $\mathcal{F}_0 \subset \Omega$.

Prove that Ω is monotone class. For increasing sequences

$$B_1 \subset B_2 \subset ... \subset B_n \subset$$

in Ω, fix $\epsilon > 0$ find $K_n \subset B_n$ compact s.t.

$$\mu(B_n) - \epsilon \leq \mu(K_n)$$

Assume that $K_n \supset K_{n-1}$, make K_{n+1}s.t.

$$\mu(B_{n+1}) - \epsilon \leq \mu(K_{n+1}) \leq \mu(B_{n+1})$$

if $K_{n+1} \not\supset K_n$ replace it with $K_{n+1} \cup K_n$

$$K_{n+1} \cup K_n \subset B_{n+1} \text{ and } \mu(K_{n+1} \cup K_n) \geq \mu(K_{n+1})$$

take limit $n \to \infty$

$$\mu(B) - \epsilon \leq \mu(K) \leq \mu(B)$$

where $B = \bigcup B_k$ and $K = \bigcup K_n$, so

$$\mu(B) = \mu(K) \implies B \in \Omega$$

For decreasing sequence $B_1 \supset B_2 \supset$ ...in Ω, fix $\epsilon > 0$ find $K_n \subset B_n$ compact s.t.

$$\mu(B_n) - \epsilon \leq \mu(K_n)$$

We cannot assume that $K_n \supset K_{n+1}$ we have to work harder. Let us do

$$\mu(B_n) - \epsilon 2^{-n} \leq \mu(K_n)$$

Take $K = \cap_{n \in \mathbb{N}} K_n$, then

$$\begin{aligned}
\mu(B) - \mu(K) &= \mu(B - K) = \mu(\cap B_n - \cap K_m) = \mu(\cap B_n \cap (\cup K_m^c)) \\
&= \mu(\cup_m((\cap_n B_n) - K_m)) \leq \mu(\cup_m(B_m - K_m)) \\
&\leq \sum \mu(B_m - K_m) \leq \sum \epsilon 2^{-n} = \epsilon
\end{aligned}$$

so $\mu(B) = \mu(K)$.

If μ is not finite, $\mathbb{R}^k = \cup_{n \in \mathbb{N}} B_n$, B_n has finite measure. Apply the previous result to $\mu_k(A) = \mu(A \cap B_k)$. If $B \in B(\mathbb{R}^k)$, $\cup_{n=1}^N B \cap B_n \nearrow B$ and to each $B \cap B_n$ we have an approximation with a compact.

2) Assume that μ is finite. $\mu(\mathbb{R}^k) = 1$

$$\begin{aligned} \mu(B) &\leq \inf\{\mu(U) : U \supset B \quad U \text{ open}\} \\ &\leq \inf\{\mu(K^c) : K^c \supset B \quad K \text{ compact}\} \end{aligned}$$

the last inequality is because the set gets smaller, not all complement of open are compact.

$$\begin{aligned} \inf\{\mu(K^c) : K^c \supset B \quad K \text{ compact}\} &= \inf\{1 - \mu(K) : K^c \supset B \text{ compact}\} \\ &= 1 - \sup\{\mu(K) : B^c \supset K,\ K \text{ compact}\} \\ &= 1 - \mu(B^c) = \mu(B) \end{aligned}$$

If μ is σ finite, $\mathbb{R}^k = \cup B_k$, B_k are Borel and bounded. If μ is LS $\implies \mu(B_k) < \infty$. If we do $\mu_k(\cdot) = \mu(\cdot \cap B_k)$, we may have a problem, for $(\cdot \cap B_k)$ may not be open, so define

$$\mu_k(\cdot) = \mu(\cdot \cap C_k)$$

where $C_k \supset B_k$, C_k open and bounded, $\mu(C_k) < \infty$. For $B \in B(\mathbb{R}^k)$ and $B \subset B_k$ apply the previous result

$$\mu_k(B) + \epsilon 2^{-k} \geq \mu_k(U_k)$$

where $U_k \supset B$ and it is open. That is

$$\mu(B \cap C_k) + \epsilon 2^{-k} \geq \mu(C_k \cap U_k)$$

$C_k \cap U_k = W_k$ is open, $\mu(B \cap C_k) = \mu(B)$.

If $A \in B(\mathbb{R}^k)$, $A \cap B_k$ is Borel and $A \cap B_k \subset B_k$. Then $\exists\, W_k$ open s.t.

$$W_k \supset A \cap B_k \text{ and } \mu(A \cap B_k) + \epsilon 2^{-k} \geq \mu(W_k)$$

$W = \cup W_k$ open

$$\mu(W) \leq \sum \mu(W_k) \quad \leq \quad \epsilon + \sum \mu(A \cap B_k) = \epsilon + \mu(A) \leq \epsilon + \mu(W)$$

therefore

$$\mu(A) = \mu(W)$$

$\square$

5.5 Product Measure

Lecture 26 -Last Lec- (5/2/13)

$(X, \mathcal{J}, \mu)$ $(Y, \mathcal{C}, \lambda)$ measure spaces, what can we say

$$(X \times Y, ?, ?)?$$

Idea: Borel is generated by interval, and product of interval is rectangle and rectangle may generate Borel in $\mathbb{R}^2$.

Definition 209. a measurable rectangle

$$A \times B : A \in \mathcal{J}, B \in \mathcal{C}$$

Let ξ the algebra generated by measurable rectangles.

Proposition 210. *ξ is made from finite disjoint unions of rectangles.*

Proof.

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$$
$$(A \times B) - (C \times D) = A \times (B - D) \cup (A - C) \times (B \cap D)$$

□

So we see

$$(X \times Y, \mathcal{J} \times \mathcal{C}, ?) \qquad \mathcal{J} \times \mathcal{C} = \sigma(\xi)$$

Technically we can slide $\mathbb{R}^2$ along x or y axis. If $E \in \mathcal{J} \times \mathcal{C}$

$$E_x = \{y; (x, y) \in E\} = P_1^{-1}(\{x\})$$
$$E_y = \{x; (x, y) \in E\} = P_2^{-1}(\{y\})$$

P is projection.

Proposition 211. *If $E \in \mathcal{J} \times \mathcal{C}$, then*

$$E_x \in \mathcal{C} \quad E_y \in \mathcal{J} \quad \forall x, y.$$

Proof. $\Omega \subset \mathcal{J} \times \mathcal{C}$ sets s.t. the statement is true, Ω contains rectangles

$$(A \times B)_x = \begin{cases} B & x \in A \\ \varnothing & x \notin A \end{cases} \in \mathcal{C}$$

If $E \in \Omega$, $(E^c)_x = (E_x)^c \implies (E^c)_x$ are measurable $\implies E^c \in \Omega$. If $E_1, E_2 \in \Omega$, since $(E_1 \cup E_2)_x = E_{1x} \cup E_{2x}$ true for arbitrary union (countable), then $E_1 \cup E_2 \in \Omega$, so Ω is σ algebra $\implies \Omega = \mathcal{J} \times \mathcal{C}$. □

Now we consider functions of 2 variables on $\mathcal{J} \times \mathcal{C}$ measurable,

$$f : X \times Y \to \mathbb{R}$$

In the same spirit, we make

$$\begin{array}{rcl} f_x : Y & \to & \mathbb{R} \\ y & \mapsto & f(x,y) \end{array} \qquad \begin{array}{rcl} f_y : X & \to & \mathbb{R} \\ x & \mapsto & f(x,y) \end{array}$$

Proposition 212. *If f is $\mathcal{J} \times \mathcal{C}$ measurable, $\mathcal{J}$, $\mathcal{C}$ complete, then f_x is $\mathcal{C}$ measurable $\forall\, x$, f_y is $\mathcal{J}$ measurable $\forall\, y$.*

Proof. $f^{-1}(V) \in \mathcal{J} \times \mathcal{C}$ if V is open. Because $f^{-1}(V) = \{(x,y); f(x,y) \in V\}$, by the previous proposition

$$f_x^{-1}(V) = \{y; f(x,y) \in V\} = \{y; f_x(y) \in V\}$$

is measurable $\forall\, x \in X$, so f_x is measurable. □

Suppose we want to find the measure of region $Q \in \mathcal{J} \times \mathcal{C}$, can we compute

$$\int_X \lambda(Q_x) d\mu(x) \text{ or } \int_Y \mu(Q_y) d\lambda(y)?$$

assuming that $\lambda(Q_x)$ is μ measurable and $\mu(Q_y)$ is λ measurable.

If they are equal, we can say the measure of $Q \in \mathcal{J} \times \mathcal{C}$ is

$$(\mu \times \lambda)(Q) := \int_Y \mu(Q_y) d\lambda(y) = \int_X \lambda(Q_x) d\mu(x) =: \int_{X \times Y} \chi_Q d(\mu \times \lambda)$$

Proposition 213. *μ, λ are σ finite, then $\lambda(Q_x)$ is measurable, $\mu(Q_y)$ is measurable and*

$$\int_X \lambda(Q_x) d\mu(x) = \int_Y \mu(Q_y) d\lambda(y) \quad \forall\, Q \in \mathcal{J} \times \mathcal{C}.$$

The σ finite assumption is very important.

Proof.

$$\lambda(Q_x) = \int_{Q_x} d\lambda = \int_Y \chi_{Q_x}(y) d\lambda(y) = \int_Y \chi_Q(x,y) d\lambda(y)$$

So we ought to show

$$\int_X \left(\int_Y \chi_Q(x,y) d\lambda(y) \right) d\mu(x) = \int_Y \left(\int_X \chi_Q(x,y) d\mu(x) \right) d\lambda(y)$$

Assume μ, λ are finite. Let $\Omega \subset \mathcal{J} \times \mathcal{C}$, $Q \in \Omega$ iff the statement is true.

Clearly rectangle $A \times B \in \Omega$, indeed

$$\lambda((A \times B)_x) = \lambda(B)\chi_A(x) \; \mu((A \times B)_y) = \mu(A)\chi_B(y)$$

Make Q_1, Q_2, ... $\in \Omega$ disjoint.

$$\chi_{\bigcup_{i=1}^n Q_i} = \sum_{i=1}^n \chi_{Q_i} \implies \bigcup_{i=1}^n Q_i \in \Omega,$$

then finite disjoint+increasing by monotone convergence $\implies \Omega$ contains countable unions. Ω is σ algebra.

Prove Ω is a monotone class.

Let $Q_n \in \Omega$, $Q_1 \subset Q_2 \subset ...$, and $Q = \bigcup_{n \in \mathbb{N}} Q_n$, then

$$\psi_n(x) = \lambda(Q_{nx}), \quad \phi_n(y) = \mu(Q_{ny})$$

are measurable and $\int_X \psi_n(x) d\mu = \int_Y \phi_n(y) d\lambda$. Also

$$\psi_n(x) \nearrow \psi(x), \quad \phi_n(y) \nearrow \phi(y)$$

are measurable.

Use monotone convergence

$$\int_X \psi(x) d\mu = \int_Y \phi(y) d\lambda$$

Since $\psi(x) = \lambda(Q_x) = \lambda(\bigcup_n Q_{nx})$,

$$\psi_n(x) = \lambda(Q_{nx}) \to \psi(x) = \lambda(Q_x)$$

and $\phi(y) = \mu(Q_y)$, therefore $Q \in \Omega$.

Take decreasing sequence $A \times B \supset Q^1 \supset Q^2 \supset ...$ and $Q^i \in \Omega$, $\mu(A) < \infty$,$\mu(B) < \infty$. Same argument show $\bigcap Q^i \in \Omega$, but use dominated convergence. Hence

$$\Omega = \mathcal{J} \times \mathcal{C}$$

Now consider μ, λ are σ finite, Let $M \subset \mathcal{J} \times \mathcal{C}$, $Q \in M$ iff the statement is true. so

$$X = \bigcup A_n, Y = \bigcup B_n, \text{ and } \mu(A_n) < \infty,\ \lambda(B_n) < \infty$$

$A_n,\ B_n$ disjoint $\implies X \times Y = \bigcup_{n,m} A_n \times B_m$. Take $Q \subset X \times Y$, let

$$Q_{nm} = Q \cap (A_n \times B_m)$$

then $M \subset \mathcal{J} \times \mathcal{C}$ $Q \in M \iff Q_{nm} \in \Omega\, \forall\, n, m.$

Take $Q^1 \supset Q^2 \supset \dots \quad Q^i \in M$, then $Q^1_{mn} \supset Q^2_{mn} \supset \dots \quad \forall\, m, n,\ Q^i_{mn} \in \Omega$, so

$$\bigcap_k Q^k_{mn} \in \Omega$$

$$\bigcap_{i \in \mathbb{N}} Q^i = \bigcup_{nm} \Big(\bigcap_{i \in \mathbb{N}} Q^i_{nm}\Big) \implies \bigcap Q^i \in M$$

showing M is monotone it contains the algebra so

$$M = \mathcal{J} \times \mathcal{C}$$

$\square$

Theorem 214. *(Tonelli) $(X \times Y, \mathcal{J} \times \mathcal{C}, \mu \times \lambda)$ σ finite $f \geq 0$, f is $\mathcal{J} \times \mathcal{C}$ measurable, then*

1)

$$\int_X \left(\int_Y f d\lambda\right) d\mu = \int_Y \left(\int_X f d\mu\right) d\lambda =: \int_{X \times Y} f d(\mu \times \lambda)$$

2)If $f \in L^1(\mu \times \lambda)$, not necessary ≥ 0, then

$$\psi(x) = \int_Y f d\lambda, \quad \phi(y) = \int_X f d\mu$$

are L^1, and f_x is $L^1(\lambda)$ a.e. wrt μ, f_y is $L^1(\mu)$ a.e. wrt λ, and

$$\int_X \left(\int_Y f d\lambda\right) d\mu = \int_Y \left(\int_X f d\mu\right) d\lambda =: \int_{X \times Y} f d(\mu \times \lambda)$$

1) function is positive, find simple functions to prove it. Unlike 2), 1) allows integral diverges

2) is called Fubini. It is totally possible

$$\int |f| d(\mu \times \lambda) < \infty$$

but the internal integral can diverge on a 0 measure set, i.e.

$$\int_X \underbrace{\int_Y f d\lambda d\mu}_{=\infty} \text{ or } \int_Y \underbrace{\int_X f d\mu d\lambda}_{=\infty}$$

Example 215. $X = [0,1]$ μ is Lebesgue, λ is counting measure. Notice counting measure over $\mathbb{R}$ is not σ finite.

$(X \times X, \mu \times \lambda)$

$$f(x,y) = \chi_{\{x=y\}} = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases}$$

compute $\int\limits_{[0,1]\times[0,1]} f(x,y) d(\mu \times \lambda)$. First do, fix y

$$\int_{[0,1]} \underbrace{\int_{[0,1]} \chi_{\{x=y\}} d\mu(x)}_{=0} d\lambda(y) = 0$$

but

$$\int_{[0,1]} \underbrace{\int_{[0,1]} \chi_{\{x=y\}} d\lambda(y)}_{=1} d\mu(x) = 1$$

If we replace $[0,1]$ by $[0,1] \cap \mathbb{Q}$, then the two integrals agree to 0.

Example 216. Let $\delta_n \nearrow 1$, $\delta_1 = 0$, and take $g_n \geq 0$ and $\mathrm{supp} g_n \subset [\delta_n, \delta_{n-1}]$. Assume by normalization

$$\int_{[0,1]} g_n d\mu = 1 = \int_{[\delta_n, \delta_{n+1}]} g_n(x) d\mu(x)$$

Let

$$f(x,y) = \sum_{n=1}^{\infty} (g_n(x) - g_{n+1}(x)) g_n(y)$$

on $[0,1] \times [0,1]$.

Consider

$$\int_{[0,1]} d\mu \int_{[0,1]} d\lambda \sum_{n=1}^{\infty} ((g_n(x) - g_{n+1}(x)) g_n(y))$$

swap integral and summation because telescopically only a few n gives non-zero,

$$\begin{aligned}
&= \int d\mu \sum_{n=1}^{\infty} \int d\lambda \left((g_n(x) - g_{n+1}(x))g_n(y)\right) = \int_{[0,1]} d\mu \sum_{n=1}^{\infty} ((g_n(x) - g_{n+1}(x)) \\
&= \sum_{m=1}^{\infty} \int_{[\delta_m,\delta_{m+1}]} d\mu \sum_{n=1}^{\infty} ((g_n(x) - g_{n+1}(x)) = \int_{[0,1]} g(x) d\mu = 1
\end{aligned}$$

But if we compute

$$\int_{[0,1]} d\lambda \int_{[0,1]} d\mu \sum_{n=1}^{\infty} \left((g_n(x) - g_{n+1}(x))g_n(y)\right)$$

$$\begin{aligned}
&= \int d\lambda \sum_{n=1}^{\infty} \int d\mu \left((g_n(x) - g_{n+1}(x))g_n(y)\right) \\
&= \int d\lambda \sum_{m=1}^{\infty} \int_{[\delta_m,\delta_{m+1}]} d\mu \sum_{n=1}^{\infty} \left((g_n(x) - g_{n+1}(x))g_n(y)\right) \\
&= \int d\lambda \sum_{m=1}^{\infty} \underbrace{\int_{[\delta_m,\delta_{m+1}]} d\mu g_m(x)}_{=1} g_m(y) - \sum_{m=2}^{\infty} \underbrace{\int_{[\delta_m,\delta_{m+1}]} d\mu g_m(x)}_{=1} g_{m-1}(y) \\
&= \int d\lambda \underbrace{\sum_{m=1}^{\infty} g_m(y) - \sum_{m=2}^{\infty} g_{m-1}(y)}_{=0} = 0
\end{aligned}$$

Fubini fails, because it's not $L^1(\mu \times \lambda)$, try

$$\iint \left| \sum_{n=1}^{\infty} \left((g_n(x) - g_{n+1}(x))g_n(y)\right) \right| d\mu(x) d\mu(y)$$

A few words about completion of product space. Is the product of two Lebesgue completion Lebesgue completion?

ANS: NO for very trivial issue.

We know one can find $E \subset [0,1] \subset \mathbb{R}$ not Lebesgue measurable, so consider $(\mathbb{R} \times \mathbb{R}, m \times m, \mu \times \mu)$ μ Lebesgue

Is this complete?

Pick $A \subset \mathbb{R}$, $\mu(A) = 0$

$$(\mu \times \mu)(A \times [0,1]) = 0$$
$$A \times E \subset A \times [0,1]$$