Math Journal

Isabelle Mills

August 29, 2025

8/31/2024

My goal for today is to work through the appendix to chapter 1 in Baby Rudin. This appendix focuses on constructing the real numbers using Dedikind cuts.

We define a cut to be a set $\alpha \subset \mathbb{Q}$ such that:

1. $\alpha \neq \emptyset$
2. If $p \in \alpha, \quad q \in \mathbb{Q}$, and $q < p$, then $q \in \alpha$.
3. If $p \in \alpha$, then $p < r$ for some $r \in \alpha$

Point 3 tells us that α doesn't have a max element. Also, point 2 directly implies the following facts:

a. If $p \in \alpha, \quad q \in \mathbb{Q}$, and $q \notin \alpha$, then $q > p$.
b. If $r \notin \alpha, \quad r, s \in \mathbb{Q}$, and $r < s$, then $s \notin \alpha$.

As a shorthand, I shall refer to the set of all cuts as R.

An example of a cut would be the set of rational numbers less than 2.

Firstly, we shall assign an ordering to R. Specifically, given any $\alpha, \beta \in R$, we say that $\alpha < \beta$ if $\alpha \subset \beta$ (a proper subset).

Here we prove that $<$ satisfies the definition of an ordering.

I. It's obvious from the definition of a proper subset that at most one of the following three things can be true: $\alpha < \beta, \quad \alpha = \beta$, and $\beta < \alpha$.

   Now let's assume that $a \not< \beta$ and $\alpha \neq \beta$. Then $\exists p \in \alpha$ such that $p \notin \beta$. But then for any $q \in \beta$, we must have by fact b. above that $q < p$. Hence $q \in \alpha$, meaning that $\beta \subset \alpha$. This proves that at least one of the following has to be true: $\alpha < \beta, \quad \alpha = \beta$, and $\beta < \alpha$.

II. If for $\alpha, \beta, \gamma \in R$ we have that $\alpha < \beta$ and $\beta < \gamma$, then clearly $\alpha < \gamma$ becuase $\alpha \subset \beta \subset \gamma$.

Now we claim that R equipped with $<$ has the least-upper-bound property.

Proof:

Let $A \subset R$ be nonempty and $\beta \in R$ be an upper bound of A. Then set $\gamma = \bigcup_{\alpha \in A} \alpha$. Firstly, we want to show that $\gamma \in R$

Since $A \neq \emptyset$, there exists $\alpha_0 \in A$. And as $\alpha_0 \neq \emptyset$ and $\alpha_0 \subseteq \gamma$ by definition, we know that $\gamma \neq \emptyset$. At the same time, we know that $\gamma \subset \beta$ since $\forall \alpha \in A$, $\alpha \subset \beta$. Hence, $\gamma \neq \mathbb{Q}$, meaning that γ satisfies property 1. of cuts.

Next, let $p \in \gamma$ and $q \in \mathbb{Q}$ such that $q < p$. We know that for some $\alpha_1 \in A$, we have that $p \in \alpha_1$. Hence by property 2. of cuts, we know that $q \in \alpha_1 \subset \gamma$, thus showing that γ satisfies property 2. of cuts.

Thirdly, by property 3. we can pick $r \in \alpha_1$ such that $p < r$ and $r \in \alpha_1 \subset \gamma$. So, γ satisfies property 3. of cuts.

With that, we've now shown that $\gamma \in R$. Clearly, γ is an upper bound of A since $\alpha \subset \gamma$ for all $\alpha \in A$. Meanwhile, consider any $\delta < \gamma$. Then $\exists s \in \gamma$ such that $s \notin \delta$. And since $s \in \gamma$, we know that $s \in \alpha$ for some $\alpha \in A$. Hence, $\delta < \alpha$, meaning that δ is not an upper bound of A. This shows that $\gamma = \sup A$.

Secondly, we want to assign $+$ and $\cdot$ operations to R so that R is an ordered field.

To start, given any $\alpha, \beta \in R$, we shall define $\alpha + \beta$ to be the set of all sums $r + s$ such that $r \in \alpha$ and $s \in \beta$.

Here we show that $\alpha + \beta \in R$.

1. Clearly, $\alpha + \beta \neq \emptyset$. Also, take $r' \notin \alpha$ and $s' \notin \beta$. Then $r' + s' > r + s$ for all $r \in \alpha$ and $s \in \beta$. Hence, $r' + s' \notin \alpha + \beta$, meaning that $\alpha + \beta \neq \mathbb{Q}$.

Now let $p \in \alpha + \beta$. Thus there exists $r \in \alpha$ and $s \in \beta$ such that $p = r + s$.

2. Suppose $q < p$. Then $q - s < r$, meaning that $q - s \in \alpha$. Hence, $q = (q - s) + s \in \alpha + \beta$.

3. Let $t \in \alpha$ so that $t > r$. Then $p = r + s < t + s$ and $t + s \in \alpha + \beta$.

Also, we shall define 0^* to be the set of all negative rational numbers. Clearly, 0^* is a cut. Furthermore, we claim that $+$ satisfies the addition requirements of a field with 0^* as its 0 element.

Commutativity and associativity of $+$ on R follows directly from the commutativity and associativity of addition on the rational numbers.

Also, for any $\alpha \in R$, $\quad \alpha + 0^* = \alpha$.

If $r \in \alpha$ and $s \in 0^*$, then $r + s < r$. Hence $r + s \in \alpha$, meaning that $\alpha + 0^* \subseteq \alpha$. Meanwhile, if $p \in \alpha$, then we can pick $r \in \alpha$ such that $r > p$. Then, $p - r \in 0^*$ and $p = r + (p - r) \in \alpha + 0^*$. So, $\alpha \subseteq \alpha + 0^*$.

Finally, given any $\alpha \in R$, let $\beta = \{p \in \mathbb{Q} \mid \exists r \in \mathbb{Q}^+ \ s.t. \ -p - r \notin \alpha\}$.

To give some intuition on this definition, firstly we want to guarentee that for all $p \in \beta$, $-p$ is greater than all elements of α. Secondly, we add the $-r$ term to guarentee that β doesn't have a maximum element.

We claim that $\beta \in R$ and $\beta + \alpha = 0^*$. Hence, we can define $-\alpha = \beta$.

To start, we'll show that $\beta \in R$:

1. For $s \notin \alpha$ and $p = -s - 1$, we have that $-p - 1 \notin \alpha$. Hence, $p \in \beta$, meaning that $\beta \neq \emptyset$. Meanwhile, if $q \in \alpha$, then $-q \notin \beta$ because there does not exist $r > 0$ such that $-(-q) - r = q - r \notin \alpha$. So $\beta \neq \mathbb{Q}$.

Now let $p \in \beta$ and pick $r > 0$ such that $-p - r \notin \alpha$.

2. Suppose $q < p$. Then $-q - r > -p - r$, meaning that $-q - r \notin \alpha$. Hence, $q \in \beta$.

3. Let $t = p + \frac{r}{2}$. Then $t > p$ and $-t - \frac{r}{2} = -p - r \notin \alpha$, meaning $t \in \beta$.

Now that we've proved $\beta \in R$, we next prove that β is the additive inverse of α. To start, suppose $r \in \alpha$ and $s \in \beta$. Then $-s \notin \alpha$, meaning that $r < -s$. So $r + s < 0$, thus showing that $\alpha + \beta \subseteq 0^*$.

As for the other inclusion, pick any $v \in 0^*$ and set $w = -\frac{v}{2}$. Then because $w > 0$, we can use the archimedean property of $\mathbb{Q}$ to say that there exists $n \in \mathbb{Z}$ such that $nw \in \alpha$ but $(n+1)w \notin \alpha$. Put $p = -(n+2)w$. Then $p \in \beta$ because $-p - w = (n+1)w \notin \alpha$. And finally, $v = nw + p \in \alpha + \beta$. Thus, $0^* \subseteq \alpha + \beta$.

9/1/2024

Based on the definition of $+$, it's also hopefully clear that for any $\alpha, \beta, \gamma \in R$ such that $\alpha < \beta$, we have that $\alpha + \gamma < \beta + \gamma$.

Next, we shall define multiplication on R. Except, first we're going to limit ourselves to the set R^+ of all cuts greater than 0^*. So, given any $\alpha, \beta \in R^+$, we shall define $\alpha\beta$ to be the set of all $p \in \mathbb{Q}$ such that $p \leq rs$ where $r \in \alpha, \quad s \in \beta, \quad r > 0$, and $s > 0$.

Here we show that $\alpha\beta \in R^+$.

1. Clearly $\alpha\beta \neq \emptyset$. Also, take any $r' \notin \alpha$ and $s' \notin \beta$. Then $r's' > rs$ for all $r \in \alpha \cap \mathbb{Q}^+$ and $s \in \beta \cap \mathbb{Q}^+$ since all four rational numbers are positive. By extension, $r's'$ is greater than all the elements (both positive and negative) of $\alpha\beta$. So, $r's' \notin \alpha\beta$, meaning that $\alpha\beta \neq \mathbb{Q}$.

Now let $p \in \alpha\beta$. Based on our definition of $\alpha\beta$, we know that the conditions of a cut trivially hold for any negative p. So, we'll assume from now on that $p > 0$. (Also note that a positive choice of p must exist because both α and β by assumption have positive elements.)

Since $p \in \alpha\beta \cap \mathbb{Q}^+$, we know there exists $r \in \alpha$ and $s \in \beta$ such that $p = rs$ and $r, s > 0$.

2. Suppose $0 < q < p$ (the case where $q \leq 0$ is trivial). Then $\frac{q}{s} < r$, meaning that $\frac{q}{s} \in \alpha$. So, $q = \frac{q}{s} \cdot s \in \alpha\beta$.

3. Let $t \in \alpha$ so that $t > r$. Then $p = rs < ts$ and $ts \in \alpha\beta$.

Also, we shall define 1^* to be the set of all rational numbers less than 1. Clearly, 1^* is a cut. And we claim that $\cdot$ satisfies the multiplication requirements of a field with 1^* as its 1 element.

As before, commutativity and associativity of $\cdot$ on R^+ follows directly from commutativity and associativity of multiplication on the rational numbers.

Next, for any $\alpha \in R^+$, we have that $\alpha 1^* = \alpha$.

It's clear that for any rational number $r \leq 0$, we have that $r \in \alpha 1^*$ and $r \in \alpha$. So we can exclusively focus on positive rational numbers.

Now suppose $r \in \alpha \cap \mathbb{Q}^+$ and $s \in 1^*$. Then $rs < r$, meaning that $rs \in \alpha$. So $\alpha 1^* \subseteq \alpha$. Meanwhile, if $p \in \alpha \cap \mathbb{Q}^+$, then we can pick $r \in \alpha$ such that $r > p$. Then $\frac{p}{r} \in 1^*$ and $p = \frac{p}{r} \cdot r \in \alpha 1^*$. So, $\alpha \subseteq \alpha 1^*$.

Thirdly, given any $\alpha \in R^+$, define:

$$\beta = \{p \in \mathbb{Q} \mid p \leq 0\} \cup \{p \in \mathbb{Q}^+ \mid \exists r \in \mathbb{Q}^+ \; s.t. \; \tfrac{1}{q} - r \notin \alpha\}$$

Here we show that $\beta \in R^+$.

1. Clearly $\beta \neq \emptyset$. Also, if $q \in \alpha$, then $\frac{1}{q} \notin \beta$. Hence, $\beta \neq \mathbb{Q}$.

Now let $p \in \beta$ and pick $r > 0$ such that $\frac{1}{p} - r \notin \alpha$. Also, assume $p > 0$ because the proof is trivial if $p \leq 0$. (The fact that $p > 0$ in β exists is trivial to show.)

2. If $q \leq 0 < p$, then trivially $q \in \beta$. Meanwhile, if $0 < q < p$, then $\frac{1}{q} - r > \frac{1}{p} - r$, meaning that $\frac{1}{q} - r \notin \alpha$. Hence, $q \notin \beta$.

3. Let $t = \frac{1}{\frac{1}{p} - \frac{r}{2}}$. Then since $\frac{1}{p} - r \notin \alpha$, we know that $\frac{1}{p} - \frac{r}{2} > 0$. Also since $\frac{1}{t} = \frac{1}{p} - \frac{r}{2} < \frac{1}{p}$, we have that $t > p$. But note that $\frac{1}{t} - \frac{r}{2} = \frac{1}{p} - r \notin \alpha$. Hence $t \notin \beta$.

We claim that $\beta\alpha = 1^*$. Hence, we can define $\frac{1}{\alpha} = \beta$.

To start, suppose $r \in \alpha \cap \mathbb{Q}^+$ and $s \in \beta \cap \mathbb{Q}^+$. Then $\frac{1}{s} \notin \alpha$, meaning that $r < \frac{1}{s}$. So $rs < 1$, thus showing that $\alpha\beta \subseteq 1^*$.

The other inclusion has a more complicated proof. Firstly, take any $v \in 1^* \cap \mathbb{Q}^+$ (the proof is trivial if $v \leq 0$). Then set $w = \frac{1}{v}$, meaning that $w > 1$. Now since $\alpha \in R^+$, we know there exists $n \in \mathbb{Z}$ such that $w^n \in \alpha$ but $w^{n+1} \notin \alpha$. Then as $w^{n+2} > w^{n+1}$, we know that $\frac{1}{w^{n+2}} \in \beta$. Hence, $v^2 = w^n \frac{1}{w^{n+2}} \in \alpha\beta$.

Now that we've shown that the square of every $v \in 1^* \cap \mathbb{Q}^+$ is also in $\alpha\beta$, we next show that there exists $z \in 1^* \cap \mathbb{Q}^+$ such that $z^2 > v$. Suppose $v = \frac{p}{q}$ where $p, q \in \mathbb{Z}^+$. Then set $z = \frac{p+q}{2q}$. Importantly, since $p < q$, we still have that $z \in 1^*$. But also note that:

$$z^2 - v = \frac{p^2+2pq+q^2}{4q^2} - \frac{4pq}{4q^2} = \frac{p^2-2pq+q^2}{4q^2} = \left(\frac{p-q}{2q}\right)^2 \geq 0$$

Thus as $v \leq z^2$ and $z^2 \in \alpha\beta$, we have that $v \in \alpha\beta$ as well. So $1^* \subseteq \alpha\beta$.

Finally, so long as $\alpha, \beta, \gamma \in R^+$, we have that $\alpha(\beta+\gamma) = \alpha\beta + \alpha\gamma$ because the rational numbers satisfy the distributive property.

Notably, in proving that $\alpha\beta \in R^+$ before, we also guarenteed that for $\alpha, \beta > 0$, we have that $\alpha\beta > 0$.

9/7/2024

Now we still need to extend our definition of multiplication from R^+ to all of R. To do this, set $\alpha 0^* = 0^* \alpha = 0^*$ and define:

$$\alpha\beta = \begin{cases} (-\alpha)(-\beta) & \text{if } \alpha < 0^*, \beta < 0^* \\ -((-\alpha)\beta) & \text{if } \alpha < 0^*, \beta > 0^* \\ -(\alpha(-\beta)) & \text{if } \alpha > 0^*, \beta < 0^* \end{cases}$$

Having done that, reproving those properties of multiplication on all of R just becomes a matter of addressing many cases and using the identity that $(-(-\alpha)) = \alpha$.

Note that that identity can be proven just from the addition properties of a field.

Because I'm bored with this construction at this point, I'm going to skip reproving those properties.

So now that we've established that R is a field, all we have left to do is to show that all numbers $r, s \in \mathbb{Q}$ are represented by cuts $r^*, s^* \in R$ such that:

- $(r+s)^* = r^* + s^*$
- $(rs)^* = r^* s^*$
- $r < s \Longleftrightarrow r^* < s^*$

Again, I'm super bored and demotivated at this point. So, I'm going to skip showing this.

With all that done, we've now shown that R satisfies all of the properties of real numbers. That concludes the proof of the existence theorem of the real numbers.

9/9/2024

Today I'm just looking at James Munkres' book *Topology*. Now while I'm done with the era of my life of taking exhaustive notes on a textbook, I still want to write down some interesting proofs. I also hope to do some exercises.

Theorem 7.8: Let A be a nonempty set. There is no injective map $f : \mathcal{P}(A) \longrightarrow A$ and there is no surjective map $g : A \longrightarrow \mathcal{P}(A)$.

In other words, the power set of a set has strictly greater cardinality.

Proof:
If such an injective f existed, then that would imply a surjective g exists. So, we just need to show that any function $g : A \longrightarrow \mathcal{P}(A)$ isn't surjective.

Let $g : A \longrightarrow \mathcal{P}(A)$ be any function and define $B = \{a \in A \mid a \in A - g(a)\}$. Clearly, $B \subseteq A$. However, B cannot be in the image of g. After all, suppose there exists $a_0 \in A$ such that $g(a_0) = B$. Then we get a contradiction because:

$$a_0 \in B \Longleftrightarrow a_0 \in A - g(a_0) \Longleftrightarrow a_0 \in A - B$$

Hence, $g(A) \neq \mathcal{P}(A)$ and we conclude that g can't be surjective. ■

Exercise 7.3: Let $X = \{0, 1\}$. Show there is a bijective correspondence between the set $\mathcal{P}(\mathbb{Z}_+)$ and the Cartesian product X^ω.

For any set $A \in \mathcal{P}(\mathbb{Z}_+)$, define $f(A)$ to be the ω-tuple $\mathbf{x}$ such that for all $i \in \mathbb{Z}^+$, $\mathbf{x}_i = 1$ if $i \in A$ and $\mathbf{x}_i = 0$ if $i \notin A$. Then clearly f is injective as $\forall A, B \in \mathcal{P}(\mathbb{Z}_+)$, $f(A) = f(B) \Longrightarrow A = B$. Also, given any $\mathbf{x} \in X^\omega$, we know that the set $A = \{i \in \mathbb{Z}_+ \mid \mathbf{x}_i = 1\}$ satisfies that $f(A) = \mathbf{x}$, meaning f is surjective.

Hence, f is a bijective function between $\mathcal{P}(\mathbb{Z}_+)$ and X^ω.

Note that this construction still works if $\mathbb{Z}_+$ is replaced with any countably infinite set.

Exercise 7.5: Determine whether the following sets are countable or not.

(f) The set F of all functions $f : \mathbb{Z}_+ \longrightarrow \{0, 1\}$ that are "eventually zero", meaning there is a positive integer N such that $f(n) = 0$ for all $n \geq N$.

F is countable. To see why, let:

$$A_n = \{f : \mathbb{Z}_+ \longrightarrow \{0, 1\} \mid \forall i \geq n, \;\; f(i) = 0\}$$

Thus each A_n is finite (with 2^n elements) and $F = \bigcup\limits_{n=1}^{\infty} A_n$.

(g) The set G of all functions $f : \mathbb{Z}_+ \longrightarrow \mathbb{Z}_+$ that are eventually 1.

G is countable. To see why, let:

$$A_n = \{f : \mathbb{Z}_+ \longrightarrow \mathbb{Z}_+ \mid \forall i \geq n, \;\; f(i) = 1\}$$

Then each A_n has a bijective correspondence with $(\mathbb{Z}_+)^n$, meaning each A_n is countable, and $G = \bigcup_{n=1}^{\infty} A_n$.

The same argument applies to all functions $f : \mathbb{Z}_+ \longrightarrow \mathbb{Z}_+$ that are eventually any constant.

(h) The set H of all functions $f : \mathbb{Z}_+ \longrightarrow \mathbb{Z}_+$ that are eventually constant.

H is countable. To see why, let A_n be the set of all functions $f : \mathbb{Z}_+ \longrightarrow \mathbb{Z}_+$ that are eventually n. Because of part g of this exercise, we know that each A_n is countable. Also, $H = \bigcup_{n=1}^{\infty} A_n$.

(i) The set I of all two-element subsets of $\mathbb{Z}_+$

(j) The set J of all finite subsets of $\mathbb{Z}_+$.

Both I and J are countably infinite. We know this because we can define surjections from $(\mathbb{Z}_+)^2$ to I and $\bigcup_{n=1}^{\infty} (\mathbb{Z}_+)^n$ to J.

(Finite cartesian products of countable sets and unions of countably many countable sets are countable.)

Exercise 7.6.a: Show that if $B \subset A$ and there is an injection $f : A \longrightarrow B$, then $|A| = |B|$.

According to the hint, we set $A_1 = A$ and $A_n = f(A_{n-1})$ for all $n > 1$. Similarly, we set $B_1 = B$ and $B_n = f(B_{n-1})$ for all $n > 1$.

We can assume A_2 is a proper subset of B_1 because if $A_2 = B_1$, then we already have that f is a bijection. Also, as f is an injection, we know that $B_2 \subset A_2$. Thus by induction, we can conclude that:

$$A_1 \supset B_1 \supset A_2 \supset B_2 \supset A_3 \supset B_3 \supset \cdots$$

Now, the textbook recommends defining $h : A \longrightarrow B$ by:

$$h(x) = \begin{cases} f(x) & \text{if } x \in A_n - B_n \text{ for any } n \in \mathbb{Z}_+ \\ x & \text{otherwise} \end{cases}$$

I want to ask a professor about this definition because it urks me. My issue with this definition of h is that I feel like it should be possible for:

$$\bigcap_{n=1}^{\infty}(A_n \cap B_n) \neq \emptyset.$$

However, we wouldn't be able to know that some x is in that intersection and thus falls into case 2 until after an infinite number of steps.

On the other hand, $S_1 = \bigcup\limits_{n=1}^{\infty}(A_n - B_n)$ is a valid definition for a set, as is $S_2 = A - S_1$. So the definition h is valid because it's saying that $h(x) = f(x)$ if $x \in S_1$ and $h(x) = x$ if $x \in S_2$.

Maybe my issue is just that I have trouble trusting the validity of a function definition if I can't actually evaluate that function myself. Although, there are lots of functions like that that I don't have any problem with. For example, given $g(x) = 0$ if x is rational and $g(x) = 1$ if x is irrational, what is $g(\pi^2)$?

Hopefully it is clear that h is in fact a valid function from A to B. Now firstly, we shall show that h is injective.

Let $x, y \in A$ such that $x \neq y$. If there are integers n and m such that $x \in A_n - B_n$ and $y \in A_m - B_m$, then $h(x) \neq h(y)$ because f is injective. Meanwhile, if no such n or m exists, then $h(x) \neq h(y)$ because $x \neq y$.

This leaves the case that there exists $n \in \mathbb{Z}_+$ such that $x \in A_n - B_n$ but for all $m \in \mathbb{Z}_+,\ \ y \notin A_m - B_m$. Then, note that $f(x) \in f(A_n - B_n)$. And since f is injective, we thus have that $f(x) \in f(A_n) - f(B_n) = A_{n+1} - B_{n+1}$. Therefore, as $y \notin A_{n+1} - B_{n+1}$, we know that $h(x) \neq y = h(y)$.

Next, we show h is surjective.

Let $y \in B$.

Suppose there exists $n \in \mathbb{Z}_+$ such that $y \in A_n - B_n$. We know that $n \neq 1$ since $y \in B$. Thus, there must exist $x \in A_{n-1}$ such that $y = f(x) \in f(A_{n-1}) = A_n$. Furthermore, this x can't be in B_{n-1} because otherwise y would be in B_n which we know isn't true. So, $x \in A_{n-1} - B_{n-1}$, meaning that $h(x) = f(x) = y$.

Meanwhile, if no such n exists, then we simply have that $h(y) = y$. Hence, $h(A) = B$.

Thus, we've shown that h is a bijection, meaning that $|A| = |B|$.

Exercise 7.7: Show that $|\{0,1\}^\omega| = |(\mathbb{Z}_+)^\omega|$.

Firstly, obviously a bijection exists between $\{0,1\}^\omega$ and $\{1,2\}^\omega$. Also, $\{1,2\}^\omega \subset (\mathbb{Z}_+)^\omega$. So, if we can construct an injective function from $(\mathbb{Z}_+)^\omega$ to $\{1,2\}^\omega$, then we can apply the result of exercise 7.6.a to prove this exercise's claim.

We shall create this injection using a diagonalization argument. Let $x \in (\mathbb{Z}_+)^\omega$. Then we define $f(x) = y \in \{1,2\}^\omega$ as follows:

$y(1) = 2$ if $x(1) = 1$. Otherwise $y(1) = 1$.

$y(2) = 2$ if $x(1) = 2$. Otherwise $y(2) = 1$.
$y(3) = 2$ if $x(2) = 1$. Otherwise $y(3) = 1$.

$y(4) = 2$ if $x(1) = 3$. Otherwise $y(4) = 1$.
$y(5) = 2$ if $x(2) = 2$. Otherwise $y(5) = 1$.
$y(6) = 2$ if $x(3) = 1$. Otherwise $y(6) = 1$.

$y(7) = 2$ if $x(1) = 4$. Otherwise $y(7) = 1$.

$$\vdots$$

Clearly f is an injection since $f(x_1) = f(x_2)$ implies that x_1 and x_2 have the same integers at all indices.

Exercise 7.6.b: (Schroeder-Bernstein theorem) If there are injections $f : A \longrightarrow C$ and $g : C \longrightarrow A$, then A and C have the same cardinality.

I did my work on paper and now it's late and I don't want to write more tonight.

9/11/2024

Since today's my day off, I'm gonna work through Munkres' textbook *Topology* some more.

Theorem 8.4 (Principle of recursive definition): Let A be a set and let a_0 be an element of A. Suppose ρ is a function assigning an element of A to each function f mapping a nonempty section of the positive integers onto A. Then there exists a unique function $h : \mathbb{Z}_+ \longrightarrow A$ such that:

$$(*) \qquad \begin{aligned} h(1) &= a_0 \\ h(i) &= \rho(h|_{\{1,\ldots,i-1\}}) \quad \text{for } i > 1. \end{aligned}$$

Proof outline:

Part 1: Given any $n \in \mathbb{Z}_+$, there exists a function $f : \{1, \ldots, n\} \longrightarrow A$ that satisfies $(*)$.

This is obvious from induction.

Part 2: Suppose that $f : \{1, \ldots, n\} \longrightarrow A$ and $g : \{1, \ldots, m\} \longrightarrow A$ both satisfy $(*)$ for all i in their respective domains. Then $f(i) = g(i)$ for all i in both domains.

Proof:

Suppose not. Let i be the smallest integer for which $f(i) \neq g(i)$.

We know $i \neq 1$ because $f(1) = a_0 = g(1)$. But then note that $f|_{\{1,\ldots,i-1\}} = g|_{\{1,\ldots,i-1\}}$. Hence:

$$f(i) = \rho(f|_{\{1,\ldots,i-1\}}) = \rho(g|_{\{1,\ldots,i-1\}}) = g(i).$$

This contradicts that i is the smallest integer for which $f(i) \neq g(i)$.

Part 3: Let $f_n : \{1, \ldots, n\} \longrightarrow A$ be the unique function satisfying $(*)$ (uniqueness was proven in part 2). Then we define:

$$h = \bigcup_{i=1}^{\infty} f_n$$

Because of part 2, we can fairly easily show that for each $i \in \mathbb{Z}_+$, there is exactly one element in h with i as it's first coordinate. Hence, the set h defines a functions from $\mathbb{Z}_+$ to A.

Also, hopefully it's clear that h satisfies $(*)$.

Axiom of choice: Given a collection $\mathcal{A}$ of disjoint nonempty sets, there exists a set C consisting of exactly one element from each element of $\mathcal{A}$.

A few notes:

1. If we restrict $\mathcal{A}$ to being a finite collection, then there is nothing controversial about this axiom. It only becomes controversial when $\mathcal{A}$ is allowed to be infinite.
2. There are multiple instances in baby Rudin where we made an infinite number of arbitrary choices. Looking at a lot of those proofs closer, I think many of them could avoid using the axiom of choice by specifying that we had to pick rational numbers in a set. However, being able to pick elements without worrying about a preexisting choice function is way easier.

   My take away from this is that not only does it make proofs cleaner to not worry about using constructed choice functions, but it's also perfectly acceptable now-a-days to use this axiom.

Plus, some really commonly used theorems require the axiom of choice to prove them. For example, the union of countably many countable sets being countable. This makes it really easy to accidentally use the axiom of choice in a proof.

Lemma 9.2: (Existence of a choice function) Given a collection $\mathcal{B}$ of nonempty sets (not necessarily disjoint), there exists a function

$$c : \mathcal{B} \longrightarrow \bigcup_{B \in \mathcal{B}} B$$

such that $c(B)$ is an element of B for each $B \in \mathcal{B}$.

Proof:
Given any set $B \in \mathcal{B}$, we define $B' = \{(B, b) \mid b \in B\}$. Because $B \neq \emptyset$, we know that $B' \neq \emptyset$ as well. Furthermore, given $B_1, B_2 \in \mathcal{B}$ if $B_1 \neq B_2$, then we have that the first element of all the pairs in B_1' are different from that of B_2'. So B_1' and B_2' are disjoint.

Now form the collection $\mathcal{C} = \{B' \mid B \in \mathcal{B}\}$. From before, we know that $\mathcal{C}$ is a collection of disjoint sets. So by the axiom of choice, there exists a set c consisting of exactly one element from each element of $\mathcal{C}$.

This set c is a subset of $\mathcal{B} \times \bigcup_{B \in \mathcal{B}} B$ which satisfies our definition of a choice function.

Hopefully it's obvious enough why c satisfies those properties.

A set A with an order relation $<$ is said to be well-ordered if every nonempty subset of A has a smallest element.

Tangent: inductiveness of $\mathbb{Z}_+$ is equivalent to the well-orderedness of $\mathbb{Z}_+$

This proof is taken from https://math.libretexts.org/ on their page for the well-ordering principle.

($\Longrightarrow$)
Suppose S is a nonempty subset of $\mathbb{Z}_+$ with no least element. Then let R be the set of lower bounds of S. Since 1 is the least element of $\mathbb{Z}_+$, we know that $1 \in R$.

Now given any $k \geq 1$, if $k \in R$, we know that $\{1, \ldots, k\}$ must be a subset of R. Also note that $R \cap S = \emptyset$ because if that wasn't true, we'd know that S has a least element. Therefore, $\{1, \ldots, k\} \cap S = \emptyset$. But then that shows that $k+1 \notin S$ since otherwise $k+1$ would be the least element of S. Furthermore, since no element of $\{1, \ldots, k\}$ is in S, we automatically have that $k+1 \in R$.

By induction, this means that $R = \mathbb{Z}_+$. Hence, we get a contradiction as S must be empty.

($\Longleftarrow$)
Let S be a subset of $\mathbb{Z}_+$ such that $1 \in S$ and $k \in S \Longrightarrow k+1 \in S$. Then suppose that $S \neq \mathbb{Z}_+$. In that case, we know that $S^{\mathsf{C}} \neq \emptyset$, and since $\mathbb{Z}_+$ is well-ordered, we know there is a least element α of S^{C}.

Because $1 \in S$, we know that $\alpha \geq 2$. But then consider that $1 \leq \alpha - 1 < \alpha$. Therefore, $\alpha - 1 \in S$, thus implying that $\alpha \in S$. This contradicts that $\alpha \in S^{\mathsf{C}}$.

From what I've heard, when defining the positive integers, one usally takes one of the two above properties as an axiom and then proves the other as a theorem. In Munkres' book, he starts with induction and proves well-orderedness.

Facts:

- If A with the order relation $<$ is well-ordered, then any subset of A is well-ordered as well with $<$ restricted to that subset.
- If A has the order relation $<_1$ and B has the order relation $<_2$ and both are well-ordered, then $A \times B$ is well-ordered with the dictionary order.
- Given any countable set A, we know there exists a bijection f from A to $\mathbb{Z}_+$. Hence, given $a, b \in A$, we can say that $a < b \Longleftrightarrow f(a) < f(b)$. Then, A is well-ordered by $<$ with the least element of any subset S of A being the element $\alpha \in A$ such that $f(\alpha)$ is the least element in $f(S)$.
- If a set A is well-ordered, then we can make a choice function $c : \mathcal{P}(A) \longrightarrow A$ using that well-ordering.

  Specifically, given any $B \subseteq A$, assign $c(B) = \beta$ where β is the least element of B.

  This is why we can pick elements of $\mathbb{Q}$ without worrying about the axiom of choice.

An important theorem (which I will hopefully prove soon) is:

The Well Ordering Theorem: If A is a set, there exists an order relation on A that is well-ordering.

Note: this theorem requires the axiom of choice to prove.

Exercise 10.5: Show that the well-ordering theorem implies the (infinite) axiom of choice.

Let $\mathcal{A}$ be a collection of disjoint sets. By the well-ordering theorem, we can pick an order relation on $\bigcup\limits_{A \in \mathcal{A}} A$ that is well-ordering.

Note that the previous sentence is carefully worded to only make use of the finite axiom of choice. Specifically, the order relation we are picking is an element of some subset of $\bigcup\limits_{A\in\mathcal{A}} A \times \bigcup\limits_{A\in\mathcal{A}} A$.

If we had instead picked a well-ordering for each $A \in \mathcal{A}$, then that would require the axiom of choice as we would be making potentially infinitely many arbitrary choices of order relations.

Now let $C = \{a \in \bigcup\limits_{A\in\mathcal{A}} A \mid \exists A \in \mathcal{A}\ s.t.\ a \in A$ and $\forall b \in A,\ \ a \leq b\}$.

Then C fulfils the properties of the set that the axiom of choice would guarentee exists.

9/14/2024

Exercise 10.1: Show that every well-ordered set has the least-upper-bound property.

Let the set A with the order relation $<$ be well-ordered. Then consider any nonempty $B \subseteq A$ and suppose there exists $\alpha \in A$ such that $b < \alpha$ for all $b \in B$.

Let $U = \{a \in A \mid \forall b \in B,\ \ b \leq a\}$. Since $\alpha \in U$, we know that $U \neq \emptyset$. So, because A is well-ordered, we know that U has a least element β. This β is by definition the least upper bound of B. So $\sup B = \beta$.

Let X be a well-ordered set. Given $\alpha \in X$, let S_α denote the set $\{x \in X \mid x < \alpha\}$. We call S_α the section of X by α.

Lemma 10.2: There exists a well-ordered set A having a largest element Ω such that S_Ω is uncountable but every other section of A is countable.

Proof:
Starting off, let B be an uncountable well-ordered set. Then let C be the well-ordered set $\{1, 2\} \times B$ with the dictionary order. Clearly, given any $b \in B$, we have that $S_{(2,b)}$ is uncountable. So the set of $c \in C$ such that S_c is uncountable is not empty.

Let Ω be the least element of C such that S_Ω is uncountable. Then define $A = S_\Omega \cup \{\Omega\}$. This is called a minimal uncountable well-ordered set.

The reason we are considering $\{1, 2\} \times B$ instead of just B is that if we were just considering B, then we wouldn't be able to guarentee that there exists $b \in B$ such that S_b is uncountable.

User MJD on https://math.stackexchange.com wrote some good intuition for why this is.

> While the set $\mathbb{Z}_+$ is countably infinite, all sections S_x of $\mathbb{Z}_+$ are finite. However, when considering $\{1,2\} \times \mathbb{Z}_+$ with the dictionary order, we have that $S_{(2,1)}$ is countably infinite. Furthermore, all sections of $S_{(2,1)}$ are finite. Thus, $S_{(2,1)}$ would be a minimal *countable* well-ordered set.

We call a set described by lemma 10.2 $\overline{S}_\Omega = S_\Omega \cup \{\Omega\}$.

Theorem 10.3: If A is a countable subset of S_Ω, then A has an upper bound in S_Ω.

Proof:
Let A be a countable subset of S_Ω. For all $a \in A$, we know that S_a is countable. Therefore, $B = \bigcup\limits_{a \in A} S_a$ is also countable, meaning that $S_\Omega - B \neq \emptyset$.

If we pick $x \in S_\Omega - B$, we must have that x is an upper bound to A because if $x < a$ for some $a \in A$, we would have that $x \in S_a \subseteq B$.

If you combine this with exercise 10.1, we know that A has a least upper bound.

Exercise 10.6: Let S_Ω be a minimal uncountable well-ordered set.

(a) Show that S_Ω has no largest element.

Suppose $\alpha \in S_\Omega$ is the largest element of S_Ω. In that case, we'd have that $S_\alpha = S_\Omega - \{\alpha\}$. However, by theorem 10.3, we know that S_α is countable. This implies that $S_\Omega = S_\alpha \cup \{\alpha\}$ must also be countable, which is a contradiction.

(b) Show that for every $\alpha \in S_\Omega$, the subset $\{x \in S_\Omega \mid \alpha < x\}$ is uncountable.

Let $\alpha \in S_\Omega$. By the law of trichotomy, we know that:

$$S_\Omega = \{x \in S_\Omega \mid x < \alpha\} \cup \{\alpha\} \cup \{x \in S_\Omega \mid \alpha < x\}.$$

Now suppose $\{x \in S_\Omega \mid \alpha < x\}$ is countable. Then as both $\{x \in S_\Omega \mid x < \alpha\}$ and $\{\alpha\}$ are countable, we have a contradiction as the three's union must also be countable. But we know S_Ω isn't.

Some definitions I've been lacking:

1. Let A be a set and suppose x, y, z are any three different elements of A.

Simple [Default] Order Relation: $(<)$	Strict Partial Order Relation: $(\prec)$
Nonreflexivity: $x \not< x$ Transitivity: $x < y$ and $y < z \Rightarrow x < z$ Comparability: $x < y$ or $y < x$ is true	Nonreflexivity: $x \not\prec x$ Transitivity: $x \prec y$ and $y \prec z \Rightarrow x \prec z$

Basically, a partial order relation is allowed to not give an order for some pairings of elements. If someone just says a set is ordered, they mean the set is simply ordered.

2. Let A and B be sets ordered by $<_A$ and $<_B$ respectively. We say that A and B have the same order type if there exists an order-preserving bijection $f : A \longrightarrow B$, meaning that $\forall a_1, a_2 \in A,\ \ a_1 <_A a_2 \Longrightarrow f(a_1) <_B f(a_2)$.

   It is trivial to show that if f is an order-preserving bijection, then f^{-1} is also an order-preserving bijection.

3. If A is an ordered set and a and b are two different elements, then consider the set $S = \{x \in A \mid a < x < b\}$. If $S = \emptyset$ we say that b is the successor of a and a is the predecessor of b.

Exercise 10.2:

(a) Show that in a well-ordered set, every element except the largest (if one exists) has an immediate successor

Let A be a well-ordered set and let α be any element in A such that there exists $\beta \in A$ for which $\alpha < \beta$. Then consider the set $S = \{x \in A \mid \alpha < x < \beta\}$. If $S = \emptyset$, then we know α has β as its successor. Meanwhile, if $S \neq \emptyset$, then since A is well-ordered, we know that A has a least element γ. Thus, the set $\{x \in A \mid \alpha < x < \gamma\} = \emptyset$ and we know that γ is the successor of α.

(b) Find a set in which every element has an immediate successor that is not well-ordered.

Consider the set $\mathbb{Z}$ of all integers using the standard ordering. Then for any $n \in \mathbb{Z}$, we know that its successor is $n + 1$. At the same time though, the set of all negative integers has no least element. So $\mathbb{Z}$ is not well-ordered by $<$.

Exercise 10.6:

(c) Let X_0 be the subset of S_Ω consisting of all elements x such that x has no immediate predecessor. Show that X_0 is uncountable.

Suppose X_0 is bounded above by some $\alpha \in S_\Omega$. Thus, there is a predecessor $x \in S_\Omega$ for any y in the set $T = \{z \in S_\Omega \mid z > \alpha\}$.

Now define a function $f : \mathbb{Z}_+ \longrightarrow T$ such that $f(1) =$ the least element of T and $f(n) =$ the successor of $f(n-1)$ for all $n > 1$. We know this function is well-defined because S_Ω has no largest element according to exercise 10.6.a. So, all elements of S_Ω and thus T have a successor by exercises 10.2.a, meaning our formula for $f(n)$ is always defined no matter what $f(n-1)$ is. Hence, the principle of recursive definition guarentees a unique f exists.

Now it's easy to show that f is injective. For suppose that given some $x, n \in \mathbb{Z}_+$ we had that $f(x) = f(x+n)$. Then that would mean that:

$$f(x) < f(x+1) < \cdots < f(x+n-1) < f(x+n) = f(x)$$

Hence we have a contradiction as $f(x) < f(x)$.

Next, we show that f is surjective. Suppose the set $R = T - f(\mathbb{Z}_+) \neq \emptyset$. Then since S_Ω and hence T is well-ordered, we know that R has a least element β. But note that β has a predecessor γ which isn't in R. More specifically, since we know that the least element of T is in $f(\mathbb{Z}_+)$, we know that γ is at least the least of element of T. So $\gamma \in T$.

Thus we conclude that $\gamma \in T - (T - f(\mathbb{Z}_+)) = f(\mathbb{Z}_+)$, meaning there exists N such that $f(N) = \gamma$. But this means that $f(N+1) = \beta$, which contradicts that β is the least element of R.

With that, we've now shown that $f : \mathbb{Z}_+ \longrightarrow T$ is a bijection, meaning that T is countable. However, this contradicts exercise 10.6.b. which asserts that T is uncountable.

Therefore, we conclude that X_0 cannot be bounded above. And by theorem 10.3, that means that X_0 can't be a countable subset of S_Ω.

Exercise 10.4:

(a) Let $\mathbb{Z}_-$ be the set of negative integers in the usual order. Show that a simply ordered set A fails to be well-ordered if and only if it contains a subset having the same order type as $\mathbb{Z}_-$.

($\Longleftarrow$)
If for some $B \subseteq A$, we have that $f : \mathbb{Z}_- \longrightarrow B$ is an order preserving bijection, then we must have that B has no least element. Hence, not all subsets of A have a least element, meaning that A is not well-ordered.

($\Longrightarrow$)
If A is not well ordered, then we know there is a set $B \subseteq A$ with no least element. Now using the axiom of choice, choose any $\beta_1 \in B$. Then for all $n > 1$, choose $\beta_n \in B_{\beta_{n-1}}$. In other words, choose $\beta_n \in B$ such that $\beta_n < \beta_{n-1}$.

Finally, define $f : \mathbb{Z}_- \longrightarrow \{\beta_n \mid n \in \mathbb{Z}_+\}$ by the rule: $f(n) = \beta_{-n}$. This f is an order preserving bijection. Thus, the set $\{\beta_n \mid n \in \mathbb{Z}_+\} \subseteq A$ has the same order type as $\mathbb{Z}_-$.

(b) Show that if A is simply ordered and every countable subset of A is well-ordered, then A is well-ordered.

It's easy to show the contrapositive of this statement.

If A is not well-ordered, then by part a. we know there exists a set $B \subseteq A$ and a function $f : \mathbb{Z}_- \longrightarrow B$ that is an order-preserving bijection. Clearly, B has no least element. Also, the function $g(n) = f(-n)$ gives a bijection from $\mathbb{Z}_+$ to B, meaning that B is countable. Hence, we have shown that B is a countable subset of A that is not well-ordered.

Let J be a well-ordered set. A subset J_0 of J is said to be inductive if for every $\alpha \in J$, we have that $(S_\alpha \subseteq J_0) \implies \alpha \in J_0$.

Exercise 10.7: (The principle of transfinite induction) If J is a well-ordered set and J_0 is an inductive subset of J, then $J_0 = J$.

Proof:
Suppose $J_0 \neq J$. That would mean the set $J - J_0$ is nonempty. So let α be the least element of $J - J_0$. We know that S_α must be disjoint to $J - J_0$, meaning that $S_\alpha \in J_0$. But then by the inductiveness of J_0, we must have that $\alpha \in J_0$. This contradicts that α is the least element of $J - J_0$.

Exercise 10.10: (Theorem) Let J and C be well-ordered sets; assume that there is no surjective function mapping a section of J onto C. Then there exists a unique function $h : J \longrightarrow C$ satisfying for each $x \in J$ the equation:

$$(*) \qquad h(x) = \text{smallest element of } C - h(S_x).$$

Proof:

(a) If h and k map sections of J or all of J into C and satisfy $(*)$ for all x in their domains, then $h(x) = k(x)$ for all x in both domains.

Proof:
Suppose not. Let y be the smallest element of the domains of h and k for which $h(y) \neq k(y)$. Then note that $\forall z \in S_y$, we must have that $h(z) = k(z)$. Thus, we get a contradiction since:

$$h(y) = \text{smallest}(C - h(S_y)) = \text{smallest}(C - k(S_y)) = k(y).$$

(b) If there exists a function $h : S_\alpha \longrightarrow C$ satisfying $(*)$, then there exists a function $k : S_\alpha \cup \{\alpha\} \longrightarrow C$ satisfying $(*)$.

Proof:
Since there is no surjective function mapping a section of J onto C, we know that $C - h(S_\alpha) \neq \emptyset$. Hence, we can define $k(x) = h(x)$ for $x < \alpha$ and $k(\alpha) =$ smallest$(C - h(S_\alpha))$.

(c) If $K \subseteq J$ and for all $\alpha \in K$ there exists $h_\alpha : S_\alpha \longrightarrow C$ satisfying $(*)$, then there exists a function $k : \bigcup\limits_{\alpha \in K} S_\alpha \longrightarrow C$ satisfying $(*)$.

Proof:
Define $k = \bigcup\limits_{\alpha \in K} h_\alpha$.

We know k is a valid function definition because part (a) guarentees that for all $\alpha_1, \alpha_2 \in K$ greater than x, we have that $h_{\alpha_1}(x) = h_{\alpha_2}(x)$. Plus, given any $x \in \bigcup\limits_{\alpha \in K} S_\alpha$, we know that there is $\alpha \in K$ such that $\forall y \in S_x,\ \ k(y) = h_\alpha(y)$. This shows that k satisfies $(*)$ at any x due to the relevant h_α satisfying $(*)$.

(d) For all $\beta \in J$, there exists a function $h_\beta : S_\beta \longrightarrow C$ satisfying $(*)$.

Proof:
Let J_0 be the set of all $\beta \in J$ for which there exists a function $h_\beta : S_\beta \longrightarrow C$ satisfying $(*)$. Our goal is to show that J_0 is inductive. That way, we can conclude by transfinite induction (exercise 10.7) that $J_0 = J$.

Pick any $\beta \in J$ and suppose $S_\beta \in J_0$.

Case 1: β has an immediate predecessor α.

Then $S_\beta = S_\alpha \cup \{\alpha\}$. So, knowing that h_α satisfying $(*)$ exists, we can use part (b) to define h_β satisfying $(*)$.

Case 2: β has no immediate predecessor.

Then $S_\beta = \bigcup\limits_{\alpha \in S_\beta} S_\alpha$.

And since we assumed that there exists $h_\alpha : S_\alpha \longrightarrow C$ satisfying $(*)$ for all $\alpha \in S_\beta$, we thus know by part (c) that there exists a function from $\bigcup\limits_{\alpha \in S_\beta} S_\alpha = S_\beta$ to C satisfying $(*)$.

Thus in both cases, we have shown that $S_\beta \in J_0$ implies that $h_\beta : S_\beta \longrightarrow C$ satisfying $(*)$ exists. Or in other words, $S_\beta \in J_0 \Longrightarrow \beta \in J_0$.

(e) Finally, we now finish proving this theorem.

Case 1: J has a max element β.

Then since we know there exists $h_\beta : S_\beta \longrightarrow C$ satisfying $(*)$, we can apply part (b) to get a function h from $J = S_\beta \cup \{\beta\}$ to C satisfying $(*)$.

Case 2: J has no max element.

Then $J = \bigcup_{\beta \in J} S_\beta$.

And since there exists $h_\beta : S_\beta \longrightarrow C$ satisfying $(*)$ for all $\beta \in J$, we can thus apply part (c) to get a function h from $J = \bigcup_{\beta \in J} S_\beta$ to C satisfying $(*)$.

9/17/2024

Theorem (The Hausdorff maximum principle): Let A be a set and let $\prec$ be a strict partial order on A. Then there exists a maximal simply ordered subset B of A.

In other words, there exists a subset B of A such that B is simply ordered by $\prec$ and no subset of A that properly contains B is simply ordered by $\prec$.

Proof:
To start out, let J be a set well-ordered by $<$ such that the elements of A are indexed in a bijective fashion by the elements of J. In other words, $A = \{a_\alpha \in A \mid \alpha \in J\}$.

Assuming the well-ordering theorem, we know that J exists. Specifically let J refer to the same set as A but equip J with the well-ordering $<$ that we know exists instead of the partial ordering $\prec$ which we equipped A.

Now our goal is to construct a function $h : J \longrightarrow \{0, 1\}$ such that $h(\alpha) = 1$ if a_α is in our maximal simply ordered subset of A and $h(\alpha) = 0$ otherwise. To do this, we rely on the **general principle of recursive definition**.

> **Theorem: (General principle of recursive definition):**
> Let J be a well-ordered set and C be any set. Given a function $\rho : \mathcal{F} \longrightarrow C$ where $\mathcal{F}$ is the set of all functions mapping sections of J into C, we have that there exists a unique functon $h : J \longrightarrow C$ satisfying that $h(\alpha) = \rho(h|_{S_\alpha})$ for all $\alpha \in J$.
>
> The proof for this is supplementary exercise 1. of this chapter. But I'm not going to do it because it's mostly identical to exercise 10.10.

Given any $\alpha \in J$ and $f : S_\alpha \longrightarrow \{0, 1\}$, define $\rho(\alpha) = 1$ if $a_\alpha \in A$ is comparable to all $a_\beta \in A$ such that $\beta \in f^{-1}(1)$ (the preimage of 1).

Note that a_α is comparable to a_β if either $a_\alpha \prec a_\beta$ or $a_\beta \prec a_\alpha$.

Then by the general principle of recursive definition, we know a unique function $h : J \longrightarrow \{0,1\}$ exists such that for all $\alpha \in J$, we have that $h(\alpha) = 1$ only when a_α is comparable to all $a_\beta \in A$ such that $\beta \in S_\alpha$ and $h(\beta) = 1$.

Let $B = \{a_\alpha \in A \mid \alpha \in J \text{ and } h(\alpha) = 1\}$. Then given any $a_\alpha, a_\beta \in B$ such that $\alpha < \beta$, we know that either $a_\alpha \prec a_\beta$ or $a_\beta \prec a_\alpha$. Hence, B is simply ordered by $\prec$. At the same time, if $a_\gamma \notin B$, then we know $h(\gamma) = 0$, meaning there exists $a_\alpha \in B$ such that $\alpha < \gamma$ and a_γ is not comparable to a_α. This shows that any set properly containing B is not simply ordered by $\prec$.

Note that the maximal simply ordered subset B is not unique. In fact, choosing a different well-ordering of J is likely to give a completely different maximal simply ordered subset.

Also, B is not empty because any set with one element is simply ordered by $\prec$.

Let A be a set and let $\prec$ be a strict partial order on A. If B is a subset of A, we say an upper bound on B is an element c of A such that for every $b \in B$, either $b = c$ or $b \prec c$. A maximal element of A is an element m of A such that for no element a of A does the relation $m \prec a$ hold.

Zorn's Lemma: Let A be a set that is strictly partially ordered. If every simply ordered subset of A has an upper bound in A, then A has a maximal element.

Proof:
By the Hausdorff maximum principle, there exists a maximal simply ordered subset B of A. Let c be an element of A that is an upperbound to B. We claim that c is a maximal element of A. For suppose there exists $d \in A$ such that $c \prec d$. We know $d \notin B$ since that would imply $d \prec c$. But by the transitivity of $\prec$, we know that $b \preceq c \prec d \Longrightarrow b \prec d$ for all $b \in B$. Hence, $B \cup \{d\}$ is simply ordered by $\prec$. This contradicts that B is a maximal simply ordered subset of A.

Exercise 11.1: If a and b are real numbers, define $a \prec b$ if $b - a$ is positive and rational.

- It's easy to show that $\prec$ is a strict partial order. After all, for all $a \in \mathbb{R}$, we have that $a - a$ is not positive. Also, if $a \prec b$ and $b \prec c$, then we know that $b - a = p$ and $c - b = q$ where $p, q \in \mathbb{Q}_+$. But then $c - a = c - b + b - a = p + q \in \mathbb{Q}_+$. So $a \prec c$.

- Clearly, given any $x \in \mathbb{R}$, the maximal simply ordered set containing x is the set $\{x + p \mid p \in \mathbb{Q}\}$.

Tangent: I never got around to writing this down last quarter. So here's a proof that assuming the axiom of choice, non-Lebesgue measurable sets exist.

Let $\mathcal{B}$ be the collection of sets of the form $S_x = [0,1] \cap \{x + p \mid p \in \mathbb{Q}\}$ where x is any real number. Obviously, all the sets in $\mathcal{B}$ are nonempty. We also claim that all the sets in $\mathcal{B}$ are disjoint. For suppose $S_x, S_y \in \mathcal{B}$ and $S_x \cap S_y \neq \emptyset$. Then fix $c \in S_x \cap S_y$ and consider any $a \in S_x$ and $b \in S_y$.

We know $c - x = p_1$, $a - x = p_2$, $c - y = q_1$, and $b - y = q_2$ where $p_1, p_2, q_1, q_2 \in \mathbb{Q}$. Thus, we have that $a - y = (a - x) + (x - c) + (c - y) = p_2 - p_1 + q_1 \in \mathbb{Q}$. Similarly, we have that $b - x = (b - y) + (y - c) + (c - x) = q_2 - q_1 + p_1 \in \mathbb{Q}$. This tells us that $a \in S_y$ and $b \in S_x$. And since this works for all $a \in S_x$ and $b \in S_y$, we thus must have that $S_x = S_y$.

Now using the axiom of choice, let V be a set containing one element from each set in $\mathcal{B}$.

To show that V is nonmeasurable, we'll reach a contradiction by supposing V is measurable. Let $q_1, q_2, \ldots$ be an enumeration of all the rational numbers in the set $[-1, 1]$. Then having defined $V + q_n = \{v + q_n \mid v \in V\}$, consider the set: $\bigcup\limits_{n \in \mathbb{Z}_+} (V + q_n)$.

Obviously, since $V \subseteq [0,1]$, we know that $\bigcup\limits_{n \in \mathbb{Z}_+} (V + q_n) \subseteq [-1, 2]$.

Also, consider any $x \in [0,1]$ and let v be the element of V which was chosen from the set $S_x \in \mathcal{B}$. Then $v - x = p$ where p is some rational number in $[-1, 1]$. So, we also know that $[0,1] \subseteq \bigcup\limits_{n \in \mathbb{Z}_+} (V + q_n)$. This means that $1 \leq \mu(\bigcup\limits_{n \in \mathbb{Z}_+} (V + q_n)) \leq 3$.

But now note that for any $n, m \in \mathbb{Z}_+$, we have that $n \neq m \Longrightarrow V + q_n \cap V + q_m = \emptyset$. To prove this, assume $V + q_n \cap V + q_m \neq \emptyset$. Thus, there would exist $v, u \in V$ such that $v + q_n = u + q_m$. In turn, we'd have that $v - u = q_m - q_n \in \mathbb{Q}$, which means that $v \in S_u$. However, this contradicts that V has only one element of S_u.

Now since μ is countably additive, we have that $\mu(\bigcup\limits_{n \in \mathbb{Z}_+} (V + q_n)) = \sum\limits_{n=1}^{\infty} \mu(V + q_n)$.

Finally, note that $\mu(V) = \mu(V + q_n)$ for all n. Thus $\sum\limits_{n=1}^{\infty} \mu(V + q_n) = \sum\limits_{n=1}^{\infty} \mu(V)$ is either 0 or ∞.

But this contradicts our earlier finding that the measure was between 1 and 3. So, we conclude that $V \notin \mathcal{M}(\mu)$. ■

Exercise 11.2:

(a) Let $\prec$ be a strict partial order on the set A. Define a (non-strict partial) relation $\preceq$ on A by letting $a \preceq b$ if either $a \prec b$ or $a = b$. Show that this relation has the following properties which are called the *partial order axioms*:

(i) $a \preceq a$ for all $a \in A$

This is true because $a = a$ for all $x \in A$.

(ii) $a \preceq b$ and $b \preceq a \implies a = b$.

Given any $a, b \in A$ such that $a \preceq b$ and $b \preceq a$, if $a \neq b$, then we'd have that $a \prec b$ and $b \prec a$. This gives a contradiction since $a \prec b \prec a \implies a \prec a$ which is not allowed.

(iii) $a \preceq b$ and $b \preceq c \implies a \preceq c$

Proving this is a matter of considering six rather trivial cases.

(b) Let P be a relation on A satisfying the three axioms above. Define a relation S on A by letting $a\,S\,b$ if $a\,P\,b$ and $a \neq b$. Show that S is a strict partial order on A.

Obviously, $a \not\!S\, a$ for all $a \in A$ since $a = a$ for all $a \in A$. Meanwhile, suppose $a\,S\,b$ and $b\,S\,c$. Then we know that $a\,P\,b$ and $b\,P\,c$, meaning that $a\,P\,c$. So we just need to show that $a \neq c$ and then we will have proven that $a\,S\,c$.

Suppose $a = c$. Then we know that $c\,P\,a$ and $a\,P\,b$, meaning that $c\,P\,b$. But then since $b\,P\,c$, we know that $b = c$. This contradicts that $b\,S\,c$.

In the next exercises we will explore some equivalent theorems to the Hausdorff maximum principle and Zorn's lemma.

Exercise 11.5: Show that Zorn's lemma implies the following:

Kuratowski's Lemma: Let $\mathcal{A}$ be a collection of sets. Suppose that for every subcollection $\mathcal{B}$ of $\mathcal{A}$ that is simply ordered by proper inclusion, the union of the elements of $\mathcal{B}$ belongs to $\mathcal{A}$. Then $\mathcal{A}$ has an element that is properly contained in no other element of $\mathcal{A}$.

To be clear, given any $A, B \in \mathcal{A}$, we defined above that $A \prec B$ if $A \subset B$. Importantly, our assumption about $\mathcal{A}$ means that every subcollection $\mathcal{B}$ of $\mathcal{A}$ that is simply ordered by $\prec$ has an upper bound in $\mathcal{A}$: $\bigcup\limits_{B \in \mathcal{B}} B$.

Thus by Zorn's lemma, we know that $\mathcal{A}$ has a maximal element C. And since there is no element $D \in \mathcal{A}$ such that $C \prec D$, we know that C is properly contained by no sets in $\mathcal{A}$.

Exercise 11.6: A collection $\mathcal{A}$ of subsets of a set X is said to be of *finite type* provided that a subset B of X belongs to $\mathcal{A}$ if and only if every finite subset of B belongs to $\mathcal{A}$. Show that the Kuratowski lemma implies the following:

Tukey's Lemma: Let $\mathcal{A}$ be a collection of sets. If $\mathcal{A}$ is of finite type, then $\mathcal{A}$ has an element that is properly contained in no other element of $\mathcal{A}$.

To start off I want to clarify that $\mathcal{A}$ being of finite types means both that:

1. For each $A \in \mathcal{A}$, every finite subset of A belongs to $\mathcal{A}$.
2. If every finite subset of a given set A belongs to $\mathcal{A}$, then A belongs to $\mathcal{A}$.

Now let $\mathcal{B}$ be any subcollection of $\mathcal{A}$ that is simply ordered by proper inclusion. Next, consider the set $S = \bigcup_{B \in \mathcal{B}} B$. We want to show that any finite subset of S is in $\mathcal{A}$.

To do this, let $n \in \mathbb{Z}_+$ and consider any subset $\{b_1, b_2, \ldots, b_n\}$ of S with n elements. Note that for each $1 \leq i \leq n$, there exists $B_i \in \mathcal{B}$ such that $b_i \in B_i$. Then since $\{B_1, B_2, \ldots B_n\}$ is a simply ordered finite set, we know that it has a maximum element B_m such that $B_i \subseteq B_m$ for all i. Hence, we have that $\{b_1, b_2, \ldots b_n\}$ is contained by some B_m in $\{B_1, B_2, \ldots, B_n\} \subseteq \mathcal{B}$. Because $\mathcal{A}$ is of finite type, this tells us that $\{b_1, b_2, \ldots b_n\} \in \mathcal{A}$.

Since we showed above that any finite subset of S is in $\mathcal{A}$, we can thus conclude because $\mathcal{A}$ is of finite type that $S \in \mathcal{A}$. And so, we have now proven the hypothesis of Kuratowski's lemma, meaning that $\mathcal{A}$ must have a set that is properly contained in other element of $\mathcal{A}$.

Exercise 11.7: Show that the Tukey lemma implies the Hausdorff maximum principle.

Let A be a set with the strict partial order $\prec$. Then let $\mathcal{A}$ be the collection of all subsets of A that are simply ordered by $\prec$. We shall show below that $\mathcal{A}$ is of finite type.

1. Suppose $B \in \mathcal{A}$. Then given any subset C of B (finite or not), we know that C is also simply ordered by $\prec$. So $C \in \mathcal{A}$.
2. Let $B \subseteq A$ and suppose every finite subset of B is in $\mathcal{A}$. Then given any two different elements $b_1, b_2 \in B$, we know that $\{b_1, b_2\} \in \mathcal{A}$, meaning that either $b_1 \prec b_2$ or $b_2 \prec b_1$. In other words, B is simply ordered by $\prec$, meaning that $B \in \mathcal{A}$.

Because $\mathcal{A}$ is of finite type, we know that $\mathcal{A}$ has an element that is properly contained in no other element of $\mathcal{A}$. Or in other words, there exists a subset of A which is simply ordered by $\prec$ and not properly contained in any other subset of A that is simply ordered by $\prec$.

9/19/2024

In the past 14 pages, we've learned a lot about the axiom of choice. All the blue arrows in the diagram to the right represent proofs we've already done. Meanwhile, the red arrows represent proofs that Munkres left to the supplementary exercises of section 1 of his book. We're gonna do those proofs now.

$$
\begin{tikzcd}
 & & \text{Non-measurable Sets} \\
\text{Axiom of Choice} \arrow[urr, Rightarrow] \arrow[d, Rightarrow, shift right] & & \\
\text{Well Ordering Theorem} \arrow[u, Rightarrow, shift right] \arrow[d, Rightarrow, shift right] & & \\
\text{Maximum Principle} \arrow[u, Rightarrow, shift right] \arrow[d, Rightarrow] & & \text{Tukey's Lemma} \arrow[ll, Rightarrow] \\
\text{Zorn's Lemma} \arrow[rr, Rightarrow] & & \text{Kuratowski's Lemma} \arrow[u, Rightarrow]
\end{tikzcd}
$$

Exercise 1: (General principle of recursive definition)

We already addressed this before. I'm skipping proving this because the proof is mostly identical to exercise 10.10. In fact, exercise 10.10 is just this exercise but with a specific $\rho : \mathcal{F} \longrightarrow C$.

Exercise 2:

(a) Let J and E be well-ordered sets and let $h : J \longrightarrow E$. Show that the following two statement are equivalent:

(i) h is order preserving and its image is E or a section of E.

(ii) $h(\alpha) = \text{smallest}(E - h(S_\alpha))$ for all $\alpha \in J$.

(i) $\Longrightarrow$ (ii):
Given any $\alpha \in J$, we know that $h(\alpha)$ must be an upper bound to $h(S_\alpha)$. Now suppose $\exists \beta \in S_{h(\alpha)}$ such that $\beta \notin h(S_\alpha)$. Because of our assumption about the image of h, we know that $\beta \in h(J)$, meaning there exists $\gamma \in J$ such that $h(\gamma) = \beta$. But because h is order-preserving, we must have that $\beta < f(\alpha) \Longrightarrow \gamma < \alpha$. This contradicts that $\beta \notin h(S_\alpha)$.

With that, we've now shown that $h(S_\alpha) = S_{h(\alpha)}$. In turn, this shows that $h(\alpha)$ is the smallest element in $E - h(S_\alpha)$.

(ii) $\Longrightarrow$ (i):
It's easy to show h is order preserving. Let $\alpha, \beta \in J$ such that $\alpha < \beta$. Then $h(S_\alpha) \subset h(S_\beta)$, meaning that $E - h(S_\beta) \subset E - h(S_\alpha)$. And since the least element of $E - h(S_\alpha)$ is not in $E - h(S_\beta)$, that means that $h(\alpha) = \text{smallest}(E - h(S_\alpha)) < \text{smallest}(E - h(S_\beta)) = h(\beta)$.

As for showing the other property of h, let $J_0 = \{\alpha \in J \mid h(S_\alpha) = S_{h(\alpha)}\}$. Now suppose that for some $\alpha \in J$, we have that $S_\alpha \subseteq J_0$. Then we can show that $\alpha \in J_0$.

Case 1: α has an immediate predecessor β.

Then $S_\alpha = S_\beta \cup \{\beta\}$, meaning that:

$$h(S_\alpha) = h(S_\beta) \cup \{h(\beta)\} = S_{h(\beta)} \cup \{h(\beta)\}.$$

Since $h(\alpha)$ is the least element of E not in $h(S_\alpha)$. We can thus say that $S_{h(\beta)} \cup \{h(\beta)\} = S_{h(\alpha)}$.

Case 2: α has no immediate predecessor.

Then we have that $h(S_\alpha) = h(\bigcup_{\beta \in S_\alpha} S_\beta) = \bigcup_{\beta \in S_\alpha} h(S_\beta) = \bigcup_{\beta \in S_\alpha} S_{h(\beta)}$.

Hence, $h(S_\alpha)$ is a section of E, and since $h(\alpha)$ is the least element not in that section, we can conclude that $h(S_\alpha) = S_{h(\alpha)}$.

By transfinite induction, we thus know that $J_0 = J$. So finally, we consider two cases.

Case 1: J has a max element α.

Then $h(J) = h(S_\alpha) \cup \{h(\alpha)\} = S_{h(\alpha)} \cup \{h(\alpha)\}$. And since $h(\alpha)$ is the least element not in $S_{h(\alpha)}$, we thus know that $h(J)$ is either a section of or the whole of E.

Case 2: J has no max element.

Then $h(J) = h(\bigcup_{\alpha \in J} S_\alpha) = \bigcup_{\alpha \in J} h(S_\alpha) = \bigcup_{\alpha \in J} S_{h(\alpha)}$.

So, $h(J)$ is either a section of or the whole of E.

(b) If E is a well-ordered set, show that no section of E has the same order type as E, nor do any two different sections of E have the same order type.

Let J be any well-ordered set. By combining part (a) of this exercise with exercise 10.10 (which is a special case of the general principle of recursive definition), we know that there is at most one order preserving map from J to E whose image is either E or a section of E. Hence, J can only have the same order type as one of either the entirety of E or one section of E.

Based on that fact, we can get an easy contradiction if we assume that the claim of part (b) is false.

9/21/2024

Unfortunately I tested positive for Covid on the two days ago. So I've been really delirious. However, right now I'm in an airport in the process of moving back out to California (great idea). And since my flight just got delayed, I feel like I might as well kill time and try to do some math.

Exercise 3: Let J and E be well-ordered sets, and suppose there is an order-preserving map $k : J \longrightarrow E$. Using exercises 1 and 2, show that J has the order type of one of either E or one section of E.

Pick any $e_0 \in E$. Then define $h : J \longrightarrow E$ by the rule:

$$h(\alpha) = \begin{cases} \text{smallest}(E - h(S_\alpha)) & \text{if } h(S_\alpha) \neq E \\ e_0 & \text{otherwise} \end{cases}$$

Note that the second case of our definition of h is just included to ensure that h is well-defined before we begin the proof in earnest. I mention that because our goal now is to show that the second case will never apply.

Let $J_0 = \{\alpha \in J \mid h(\alpha) \leq k(\alpha)\}$. Then suppose that for some $\alpha \in J$, we have that $S_\alpha \subseteq J_0$. Because k is order preserving, we know that $k(\alpha) > k(\beta) \geq h(\beta)$ for all $\beta \in S_\alpha$. Hence, $k(\alpha) \notin h(S_\alpha)$, meaning that $h(S_\alpha) \neq E$. So, we conclude that $h(\alpha) = \text{smallest}(E - h(S_\alpha))$. And since $k(\alpha) \in E - h(S_\alpha)$, we thus know that $h(\alpha) \leq k(\alpha)$

Therefore, $\alpha \in J_0$. By transfinite induction, this proves that $J = J_0$. The reason this is relevant is that we can now say that $k(\alpha)$ is never in $h(S_\alpha)$, meaning that $E - h(S_\alpha) \neq \emptyset$. So $h(\alpha)$, will never be determined by the second case of our definition above.

By exercise 2, we know that $h : J \longrightarrow E$ is the unique order-preserving map whose image is either E or a section of E. Thus, J has the same order type as exactly one of either the entirety of E or one section of E.

Exercise 4: Use exercises 1-3 to prove the following:

(a) If A and B are well-ordered sets, then exactly one of the following three conditions holds: A and B have the same order type, A has the order type of a section of B, or B has the order type of a section of A.

To start off, it's relatively easy to show that at most one of the above three cases is true. After all, A having the same order type as B as well as a section of B contradicts exercise 2. Similarly B having the same order type as A as well as a section of A contradicts exercise 2.

Meanwhile, to find a contradiction if A has the order type of S_β and B has the order type of S_α where $\alpha \in A$ and $\beta \in B$, let $h : A \longrightarrow S_\alpha$ be the function defined by the rule $h(a) = g(f(a))$ where f is the order-preserving bijection from A to S_β and g is the order-preserving bijection from B to S_α.

Then given any $a, b \in A$, we know that:

$$a < b \Rightarrow f(a) < f(b) \Rightarrow h(a) = g(f(a)) < g(f(b)) = h(b).$$

Hence, h is an order preserving map from A to S_α. This gives us a contradiction since exercise 3 would then imply that A has the same order type as either S_α or a section of S_α (which would still be a section of A).

Now, what's left to show is that at least one of the three above cases must be true. Unfortunately, the hinted route for showing this uses an exercise I didn't do. And right now I really don't want to do that exercise. So I'm just going to write out the thing I was supposed to have proven earlier.

Exercise 10.8.a:

Let A_1 and A_2 be disjoint sets well-ordered by $<_1$ and $<_2$ respectively. Then define an order relation on $A_1 \cup A_2$ by letting $a < b$ either if $a, b \in A_1$ and $a <_1 b$, or if $a, b \in A_2$ and $a <_2 b$, or if $a \in A_1$ and $b \in A_2$. This is a well-ordering of $A_1 \cup A_2$.

Let $A' = \{A\} \times A$ and let $B' = \{B\} \times B$. That way, so long as $A \neq B$, we know that A' and B' are disjoint. (The case where $A = B$ is trivial.)

It's hopefully obvious that the well-orderings of A and B can be used to well-order A' and B'. For A', define $(A, a_1) <_{A'} (A, a_2)$ if $a_1 <_A a_2$. Similarly, define the analogous ordering for B'. Clearly, A and A' have the same order type, as do B and B'. Also, given any $\alpha \in A$ and $\beta \in B$, S_α and $S_{(A,\alpha)}$ have the same order type, as do S_β and $S_{(B,\beta)}$

Next, define a well-ordering on $A' \cup B'$ by letting $a' < b'$ if either $a', b' \in A'$ and $a' <_{A'} b'$, or if $a', b' \in B'$ and $a' <_{B'} b'$, or if $a' \in A'$ and $b' \in B'$.

Note that the inclusion function from B' to $A' \cup B'$ is an order-preserving map. Thus, by exercise 3, we know that B' has the order type of one of either $A' \cup B'$ or one section of $A' \cup B'$.

Case 1: B' has the order type of a section S_α of $A' \cup B'$.

If $\alpha \in A'$, then B' has the order type of a section of A', meaning B has the order type of a section of A.

If α is the first element of B, then B' has the same order type as A', meaning B has the same order type as A.

If $\alpha \in B'$, then there exists an order preserving bijection from B' to $A' \cup \{b \in B' \mid b <_{B'} \alpha\}$. So let f be the inverse of that bijection but with it's domain restricted to just A'. Since f is also an order-preserving map, we know by exercise 3 that A' has the order type of either B' or a section of B'. This would mean that A has the order type of either B or a section of B.

Case 2: B' has the order type of $A' \cup B'$.

Let f be the inverse of the order preserving bijection from B' to A', except with it's inverse restricted to just A'. Since f is also an order-preserving map, we know by exercise 3 that A' has the order type of either B' or a section of B'. This would mean that A has the order type of either B or a section of B.

With that, we've now shown that at least one of the three cases posed by the exercise will always be true.

(b) Suppose that A and B are well-ordered sets that are uncountable such that every section of A and of B is countable. Show that A and B have the same order type.

If A did not have the same order type as B, then by part (a) of this exercise we would know that either A has the order type of a section of B or B has the order type of a section of A. However, that would suggest the existence of a bijection between a countable set and an uncountable set, which by definition is not possible.

9/23/2024

Exercise 5: Let X be any set and let $\mathcal{A}$ be the collection of all pairs $(A, <)$ where A is a subset of X and $<$ is a well-ordering of A. Define:

$$(A, <) \prec (A', <')$$

if $(A, <)$ equals a section of $(A', <')$.

In other words, $A = S_\alpha = \{a \in A' \mid a <' \alpha\}$ where $\alpha \in A'$, and $<$ is the order relation $<'$ restricted to A.

(a) Show that $\prec$ is a strict partial order on $\mathcal{A}$.

Clearly no A is a section of itself. So $(A, <) \not\prec (A, <)$.

Also if $(A, <_A) \prec (B, <_B) \prec (C, <_C)$, then we know that A is a section of a section of C (which is still a section). Plus, $<_A$ is just $<_C$ restricted to $<_A$. Hence, $(A, <) \prec (C, <_C)$.

(b) Let $\mathcal{B}$ be a subcollection of $\mathcal{A}$ that is simply ordered by $\prec$. Define B' to be the union of the sets B for all $(B,<) \in \mathcal{B}$, and define $<'$ to be the union of the relations $<$ for all $(B,<) \in \mathcal{B}$. Show that $(B',<')$ is a well-ordered set.

To start, let's quickly double check that $<'$ is a valid order relation on B'.

(i) Given any $b \in B'$, if $b \in B$ for any $(B,<) \in \mathcal{B}$, then we know that $(b,b) \notin <$. So $(b,b) \notin <'$.

(ii) Suppose $a,b \in B'$ such that $(a,b) \notin <'$. Then for all $(B,<) \in \mathcal{B}$ such that $a,b \in B$, we know that $(a,b) \notin <$, meaning that $(b,a) \in <$. So $(b,a) \in <'$.

(iii) Given $a,b,c \in B'$, suppose $a <' b <' c$. Then there exists $(B_1,<_1)$ and $(B_2,<_2)$ in $\mathcal{B}$ such that $(a,b) \in <_1$ and $(b,c) \in <_2$. Now by how we defined $\mathcal{B}$, we know that either $<_1 \subset <_2$ or $<_2 \subset <_1$. Thus, we know $(a,b),(b,c) \in \{<_i\}$ for some $i \in \{1,2\}$. Hence, $(a,c) \in <_i$, meaning that $(a,c) \in <'$.

Next, we show that B' is well-ordered by $<'$.

Let $S \subseteq B'$ be nonempty and pick any element β in S. Then we know there exists $(B_1,<_1) \in \mathcal{B}$ such that $\beta \in B_1$. Also, B_1 is well-ordered by $<$. So let α be the least element (using $<_1$) of $B_1 \cap S$.

We claim that α is the least element (using $<'$) of S. To prove this, suppose there exists $c \in S$ such that $c <' a$. Then we know $(c,\alpha) \in <_2$ for some $(B_2,<_2) \in \mathcal{B}$. Importantly, $(B_1,<_1) \neq (B_2,<_2)$ since otherwise we'd have chosen α differently. So one must be a section of the other.

- If $(B_2,<_2)$ is a section of $(B_1,<_1)$, then we know that $<_2 \subset <_1$ and $c \in B_1 \cap S$. But this contradicts how we chose α.
- If $(B_1,<_1)$ is a section of $(B_2,<_2)$, then we know there exists $\gamma \in B_2$ such that $B_1 = S_\gamma \subseteq B_2$. If $c <_2 \gamma$, then we know that $c \in B_1$ and thus $B_1 \cap S$. This contradicts how we chose α. So we must have that $\gamma <_2 c$. But then this also gives us a contradiction as $\alpha <_2 \gamma <_2 c \Longrightarrow \alpha <_2 c$, meaning that $\alpha <' c$.

(c) [Not in the book...] Given any $\mathcal{B}$ from part (b) of this problem and defining $(B',<')$ as before, we have that $(B,<) \preceq (B',<')$ for all $(B,<) \in \mathcal{B}$.

Consider any $(B_1,<_1) \in \mathcal{B}$. If $B_1 \neq B'$, then we know there exists $\alpha \in B' - B_1$, thus meaning there exists $(B_2,<_2) \in \mathcal{B}$ such that $\alpha \in B_2$. Since $B_2 \not\prec B_1$, we know that $B_1 \prec B_2$, meaning that $B_1 = S_\beta \subseteq B_2$ for some $\beta \in B_2$.

Now we know that $\{b \in B' \mid b <' \beta\} \subseteq \{b \in B_2 \mid b <_2 \beta\}$. For suppose there exists a in the former set but not the latter set. Then there must exist $(B_3, <_3) \in \mathcal{B}$ such that $(a, b) \in <_3$.

If $a \in B_2$, then we'd have that $(b, a) \in <_2$. But that would imply that (a, b) and (b, a) are in $<'$ which we know isn't possible. So we know that $B_3 \not\subseteq B_2$.

Since $\mathcal{B}$ is simply ordered by $\prec$ and we can't have that $B_3 \prec B_2$, we know that $B_2 \prec B_3$. So $B_2 = S_\gamma$ where $\gamma \in S_3$. Now $a <_3 \gamma$ would contradict that $a \notin B_2$. So we must have that $\gamma <_3 a$. However, we also must have that $b <_3 \gamma$, which contradicts that $a <_3 b$.

Hence, we've shown that $(B_1, <_1) \neq (B', <')$ implies that $(B_1, <_1) \prec (B', <')$.

Exercise 6: Use exercise 5 to prove that the maximum principle implies the well-ordering theorem.

Let X be any set and construct $\mathcal{A}$ and $\prec$ as before in exercise 5. By the maximal principle, we know there exists $\mathcal{B} \subseteq \mathcal{A}$ such that $\mathcal{B}$ is simply ordered by $\prec$ and no proper superset of $\mathcal{B}$ is simply ordered by $\prec$.

Next, construct B' and $<'$ as in exercise 5.b. We claim that $B' = X$. To see this, suppose there exists $c \in X - B'$. Then $B' \cup \{c\}$ is well-ordered by the order relation: $<' \cup \{(b, c) \mid b \in B'\}$. Hence $(B' \cup \{c\}, <' \cup \{(b, c) \mid b \in B'\}) \in \mathcal{A}$.

At the same time, note that $(B', <') \prec (B' \cup \{c\}, <' \cup \{(b, c) \mid b \in B'\})$. And since we have that $(B, <) \preceq (B', <')$ for all $(B, <) \in \mathcal{B}$, we thus know that for any $(B, <) \in \mathcal{B}$:

$$(B, <) \prec (B' \cup \{c\}, <' \cup \{(b, c) \mid b \in B'\}).$$

This tells us both that $(B' \cup \{c\}, <' \cup \{(b, c) \mid b \in B'\}) \notin \mathcal{B}$ and that $(B' \cup \{c\}, <' \cup \{(b, c) \mid b \in B'\})$ is comparable with all elements of $\mathcal{B}$. But that contradicts that $\mathcal{B}$ is a maximal simply ordered subset of $\mathcal{A}$.

So we must have that $B' = X$. And thus by exercise 5.b, we know that a well-ordering of X exists.

9/24/2024

Exercise 7: Use exercises 1-5 to prove that the choice axiom implies the well-ordering theorem.

Let X be a set and c be a fixed choice function for the nonempty subsets of X. If T is a subset of X and $<$ is a relation on T, we say that $(T, <)$ is a tower in X if $<$ is a well-ordering of T and if for each $x \in T$, $x = c(X - S_x(T))$ where $S_x(T)$ is the section of T by x.

Well, shit. I wish I was given that notation for specifying which set I was taking a section of before I did exercise 2. $h(S_x(J)) = S_{h(x)}(E)$ is a lot clearer notation than just $h(S_x) = S_{h(x)}$

(a) Let $(T_1, <_1)$ and $(T_2, <_2)$ be two towers in X. Show that either these two ordered sets are the same or one equals a section of the other.

By applying exercise 4 and switching indices if necessary, we know that T_1 has the order type of one of either T_2 or one section of T_2. In other words, there exists an order preserving map $h : T_1 \longrightarrow T_2$ such that $h(T_1)$ equals T_2 or a section of T_2.

Now we assert that given any $x \in T_1$, $h(x) = x$. To prove this, first note that because of transfinite induction, we can assume that $h(x) = x$ for all x in $S_x(T_1)$. This means that we can assume $h(S_x(T_1)) = S_x(T_1)$. Also, as part of doing exercise 2, we proved that h must satisfy that $h(S_x(T_1)) = S_{h(x)}(T_2)$. Hence, $S_x(T_1) = S_{h(x)}(T_2)$. This let's us conclude that:

$$x = c(X - S_x(T_1)) = c(X - S_{h(x)}(T_2)) = h(x).$$

With that we now know that $h(T_1) = T_1$. So T_1 equals either T_2 or a section of T_2.

(b) If $(T, <)$ is a tower in X and $T \neq X$, then there is a tower in X of which $(T, <)$ is a section.

Since $T \neq X$, let $y = c(X - T)$. Then define $T' = T \cup \{y\}$ and $<' = < \cup \{(x, y) \mid x \in T\}$. Clearly, $(T', <')$ is a tower which contains $(T, <)$ as a section.

Clearly T' is well-ordered by $<'$.

Also, if $x \in T' - \{y\}$, then we have that $c(X - S_x(T')) = c(X - S_x(T)) = x$. Plus, we know that $c(X - S_y(T')) = c(X - T) = y$.

(c) Let $\{(T_k, <_k) \mid k \in K\}$ be the collection of all towers in X. Then define:

$$T = \bigcup_{k \in K} T_k \text{ and } < = \bigcup_{k \in K} <_k.$$

Show that $(T, <)$ is a tower in X. Conclude that $T = X$.

If we define $\mathcal{A}$ and $\prec$ from X as we did in exercise 5, we can see from part (a) of this problem that $\{(T_k, <_k) \mid k \in K\}$ is a subset of $\mathcal{A}$ that is simply ordered by $\prec$. Thus, from part (b) of exercise 5, we know that T is well-ordered by $<$.

To prove that T is a tower, consider any $y \in T$. Then we know there exists $k \in K$ such that $y \in T_k$. Furthermore, we know that $y = c(X - S_y(T_k))$. By, part (c) of exercise 5, we know that T_k is either a section of T or all of T. Hence, $S_y(T) = S_y(T_k)$. And thus we have that $y = c(X - S_y(T))$.

Now that we have shown $(T, <)$ is a tower in X, we get an easy contradiction if $T \neq X$. This is because T must contain all towers, but T not equalling X would imply the existence of a tower not contained by T due to part (b) of this exercise.

And since $T = X$, we thus have that $<$ is a well-ordering of X. ■

I'm gonna skip doing exercise 8 of the supplementary exercise. Basically it shows that you can construct a well-ordered set with higher cardinality than an arbitrary well-ordered set, all without using the axiom of choice. Also, while that does mean we can construct a minimal uncountable well-ordered set without using the axiom of choice, theorem 10.3 requires the axiom of choice to prove. So almost nothing we discovered about a minimal uncountable well-ordered set can be proven without the axiom of choice.

9/25/2024

I'm gonna try to cram as much topology as I can today before class starts tomorrow. After all, I suspect and fear that a bunch of this will be necessary at some point in 240. As before, I'm shamelessly ripping off James Munkres' book.

A Topology on a set X is a collection $\mathcal{T}$ of subsets of X having the properties:

1. $\emptyset$ and X are in $\mathcal{T}$.
2. The union of the elements of any subcollection of $\mathcal{T}$ is in $\mathcal{T}$.
3. The intersection of the elements of any finite subcollection of $\mathcal{T}$ is in $\mathcal{T}$.

Technically, a topological space is an ordered pair $(X, \mathcal{T})$ consisting of a set X and a topology $\mathcal{T}$ on X. But when no confusion will arise, we usually omit mentioning $\mathcal{T}$ and just call X a topological space.

Given a topological space $(X, \mathcal{T})$, we say that a subset U of X is an open set if $U \in \mathcal{T}$.

Suppose $\mathcal{T}$ and $\mathcal{T}'$ are topologies on X Such that $\mathcal{T} \subseteq \mathcal{T}'$. Then we say $\mathcal{T}'$ is finer or larger than $\mathcal{T}$ Also, we say $\mathcal{T}$ is coarser or smaller than $\mathcal{T}'$. And we say both are comparable with each other.

If $\mathcal{T}$ is properly contained by $\mathcal{T}'$, then we add the word *strictly* before those adjectives.

If X is a set, a basis for a topology on X is a collection $\mathcal{B}$ of subsets of X (called basis elements) such that:

1. For each $x \in X$, there is at least one basis element B containing x.
2. If x belongs to the intersection of two basis elements B_1 and B_2, then there is a basis element B_3 containing x such that $B_3 \subseteq B_1 \cap B_2$.

If $\mathcal{B}$ satisfies these two conditions, then we define the topology $\mathcal{T}$ generated by $\mathcal{B}$ as follows:

$U \subseteq X$ is open if for each $x \in U$, there is a basis element $B \in \mathcal{B}$ such that $x \in B$ and $B \subseteq U$.

Proof that the $\mathcal{T}$ generated by $\mathcal{B}$ is a topology:

We fairly trivially have that $\emptyset$ and X are included in $\mathcal{T}$.

Let $\{U_\alpha\}_{\alpha \in J}$ be an indexed family of elements of $\mathcal{T}$ and define $U = \bigcup_{\alpha \in J} U_\alpha$. Given any $x \in U$, we know there exists $\alpha \in J$ such that $x \in U_\alpha$. And since U_α is open, there exists $B \in \mathcal{B}$ such that $x \in B$ and $B \subseteq U_\alpha \subseteq U$. So, we conclude that U is also open.

Finally, we shall prove by induction that given $U_1, \ldots U_n \in \mathcal{T}$, we have that $U_1 \cap \ldots \cap U_n \in \mathcal{T}$.

Firstly, consider any $U_1, U_2 \in \mathcal{T}$. Then, given any $x \in U_1 \cap U_2$, choose basis elements $B_1, B_2 \in \mathcal{B}$ such that $x \in B_1 \subseteq U_1$ and $x \in B_2 \subseteq U_2$. Since $x \in B_1 \cap B_2$, we know there is a basis element $B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \cap B_2$. Then $x \in B_3 \subseteq U$.

With that, we've now shown that the intersection of any two elements of $\mathcal{T}$ is also in $\mathcal{T}$. So, we can proceed by induction.

Suppose for $i < n$ that $(U_1 \cap \ldots \cap U_i) \in \mathcal{T}$. Then we know that $(U_1 \cap \ldots \cap U_i) \cap U_{i+1} \in \mathcal{T}$.

Lemma 13.1: Let X be a set and $\mathcal{B}$ be a basis for a topology $\mathcal{T}$ on X. Then $\mathcal{T}$ equals the collection of all unions of elements of $\mathcal{B}$.

Proof:

Let $\mathcal{T}'$ be the collection of all unions of elements of $\mathcal{B}$.

Since every $B \in \mathcal{B}$ is an element of $\mathcal{T}$, we trivially have that $\mathcal{T}' \subseteq \mathcal{T}$. Meanwhile, given any $U \in \mathcal{T}$, choose for each $x \in U$ an element B_x of $\mathcal{B}$ such that $x \in B_x \subseteq U$. Then $U = \bigcup_{x \in U} B_x$, meaning $U \in \mathcal{T}'$.

(Axiom of Choice usage alert!!)

Lemma 13.2: Let X be a topological space. Suppose that $\mathcal{C}$ is a collection of open sets of X such that for each open set U of X and each $x \in U$, there is an element $C \in \mathcal{C}$ such that $x \in C \subseteq U$. Then $\mathcal{C}$ is a basis for the topology of X.

Proof:

Firstly, we need to show that $\mathcal{C}$ is a basis.

Since X is an open set, we know by hypothesis that for all $x \in X$, there is $C \in \mathcal{C}$ such that $x \in C$. As for the second condition of a basis, suppose $x \in C_1 \cap C_2$ where $C_1, C_2 \in \mathcal{C}$. Since C_1 and C_2 are open, we know that $C_1 \cap C_2$ is open. So by hypothesis, there is $C_3 \in \mathcal{C}$ such that $x \in C_3 \subseteq (C_1 \cap C_2)$.

Secondly, we need to show that $\mathcal{C}$ is a basis for the topology of X.

Let $\mathcal{T}$ be the collection of open sets of X, and let $\mathcal{T}'$ be the topology generated by $\mathcal{C}$. Firstly, if $U \in \mathcal{T}$ and $x \in U$, there is by hypothesis $C \in \mathcal{C}$ such that $x \in C$ and $C \subseteq U$. So $U \subseteq \mathcal{T}'$. Meanwhile, if $W \in \mathcal{T}'$, then W equals a union of elements of $\mathcal{C}$ by lemma 13.1. Since each element of $\mathcal{C}$ is in $\mathcal{T}$, we know W is the union of elements of $\mathcal{T}$, meaning $W \in \mathcal{T}$. So, we've shown that $\mathcal{T} \subseteq \mathcal{T}' \subseteq \mathcal{T}$.

Lemma 13.3: Let $\mathcal{B}$ and $\mathcal{B}'$ be bases for the topologies $\mathcal{T}$ and $\mathcal{T}'$ respectively on X. Then $\mathcal{T}'$ is finer than $\mathcal{T}$ if and only if for each $x \in X$ and each basis element $B \in \mathcal{B}$ containing x, there is a basis element $B' \in \mathcal{B}'$ such that $x \in B' \subseteq B$.

Proof:

($\Longrightarrow$) Let $x \in X$ and $B \in \mathcal{B}$ such that $x \in B$. Since $B \in \mathcal{T}$ and we are assuming $\mathcal{T} \subseteq \mathcal{T}'$, we know that $B \in \mathcal{T}'$. Then since $\mathcal{B}'$ generated $\mathcal{T}'$, we know there is $B' \in \mathcal{B}'$ such that $x \in B' \subseteq B$.

($\Longleftarrow$)

Given an element U of $\mathcal{T}$, we need to show that $U \in \mathcal{T}'$. To do this, consider any $x \in U$. Since $\mathcal{B}$ generates $\mathcal{T}$, there is an element $B \in \mathcal{B}$ such that $x \in B \subseteq U$. Now by hypothesis, there exists $B' \in \mathcal{B}'$ such that $x \in B' \subseteq B$. So $x \in B' \subseteq U$. Hence, $U \in \mathcal{T}'$.

If $\mathcal{B}$ is the collection of all open intervals $(a,b) = \{x \in \mathbb{R} \mid a < x < b\}$ in the real line, then we call the topology generated by $\mathcal{B}$ the standard topology on the real line.

We assume $\mathbb{R}$ has this topology unless stated otherwise.

If $\mathcal{B}'$ is the collection of all intervals $[a,b)$ of the real line, we call the topology generated by $\mathcal{B}'$ the lower limit topology.

When $\mathbb{R}$ has this topology, we denote it $\mathbb{R}_l$.

Letting $K = \{\frac{1}{n} \mid n \in \mathbb{Z}_+\}$, if $\mathcal{B}''$ is the collection of all intervals (a,b) of the real line along with all sets of the form $(a,b) - K$, then we call the topology generated by $\mathcal{B}''$ the K-topology on the real line.

When $\mathbb{R}$ has this topology, we denote it $\mathbb{R}_K$.

Lemma 13.4: The topologies of $\mathbb{R}_l$ and $\mathbb{R}_K$ are strictly finer than the standard topology on $\mathbb{R}$. But, they aren't comparable with one another.

Proof:

Let $\mathcal{T}, \mathcal{T}', \mathcal{T}''$ be the topologies of $\mathbb{R}, \mathbb{R}_l, \mathbb{R}_K$ respectively.

Given any $(a,b) \in \mathcal{B}$ and $x \in (a,b)$, we know that $[x,b) \in \mathcal{B}'$ and that $x \in [x,b) \subseteq (a,b)$. So by lemma 13.3, $\mathcal{T} \subseteq \mathcal{T}'$. On the other hand, for any $[x,b) \in \mathcal{B}'$, there is no set $(a,b) \in \mathbb{B}$ such that $x \in (a,b) \subseteq [x,b)$. So $\mathcal{T}' \not\subseteq \mathcal{T}$. Hence, $\mathcal{T}'$ is strictly finer than $\mathcal{T}$.

Also, given any $(a,b) \in \mathcal{B}$, we also know that $(a,b) \in \mathcal{B}''$. So $\mathcal{T} \subseteq \mathcal{T}''$. On the other hand, given $(-1,1) - K \in \mathcal{B}''$, we know there is no interval $(a,b) \in \mathcal{B}$ such that $0 \in (a,b) \subseteq (-1,1) - K$. So by lemma 13.3, we know that $\mathcal{T}'' \not\subseteq \mathcal{T}$. Hence, $\mathcal{T}''$ is strictly finer than $\mathcal{T}$.

Finally, we show $\mathcal{T}'$ and $\mathcal{T}''$ aren't comparable. Firstly, given the set $(-1,1) - K$ in $\mathcal{B}''$, there is no set $[a,b) \in \mathcal{B}'$ such that $0 \in [a,b) \subseteq (-1,1) - K$. After all, for any $b > 0$, we can use the archimedean property to find $\frac{1}{n} < b$. Secondly, given the set $[0,1) \in \mathcal{B}'$, no set of the form (a,b) can satisfy that $0 \in (a,b) \subseteq [0,1)$. Similarly, no set of the form $(a,b) - K$ can satisfy that $0 \in (a-b) - K \subseteq [0,1)$. So neither $\mathcal{T}' \subseteq \mathcal{T}''$ nor $\mathcal{T}'' \subseteq \mathcal{T}'$.

A subbasis $\mathcal{S}$ for a topology on X is a collection of subsets of X whose union equals S. The topology generated by the subbasis $\mathcal{S}$ is defined to be the collection $\mathcal{T}$ of all unions of finite intersections of elements of $\mathcal{S}$.

Proof that the $\mathcal{T}$ generated by $\mathcal{S}$ is a topology:

It suffices to show that the collection $\mathcal{B}$ of all finite intersections of elements of $\mathcal{S}$ is a basis. The first condition of a basis is trivially true for $\mathcal{B}$ since the union of the elements of $\mathcal{S}$ is all of X and $\mathcal{S} \subseteq \mathcal{B}$.

As for the second condition of a basis, given any $(S_1 \cap \ldots \cap S_n), (S'_1 \cap \ldots \cap S'_m) \in \mathcal{B}$, we know that $(S_1 \cap \ldots \cap S_n) \cap (S'_1 \cap \ldots \cap S'_m)$ is a finite intersection of elements of $\mathcal{S}$ and thus an element in $\mathcal{B}$. Thus, the condition easily follows.

9/26/2024

Well, it looks like I'll be able to survive 240A with the topology information I've learned so far. However, it doesn't look like I'll be able to survive 240B with what I know right now. So, I've got to study more of this. But if needed for 188 this quarter, I'll take a break to study algebra.

Exercise 13.3 Show that $\mathcal{T} = \{U \subseteq X \mid X - U$ is countable or all of $X\}$ is a topology on X.

Clearly $\emptyset, X \in \mathcal{T}$ since $|X - X| = 0$ and $X - \emptyset = X$.

Suppose $\{U_\alpha\}_{\alpha \in A}$ is a collection of sets in $\mathcal{T}$. Then $X - \bigcup\limits_{\alpha \in A} U_\alpha = \bigcap\limits_{\alpha \in A} (X - U_\alpha)$ is countable since it's a subset of a countable set.

Hence, $\bigcup\limits_{\alpha \in A} U_\alpha \in \mathcal{T}$.

Finally, consider any $\{U_1, \ldots, U_n\}$ in $\mathcal{T}$. Then $X - \bigcap\limits_{k=1}^{n} U_k = \bigcup\limits_{k=1}^{n} (X - U_k)$ is countable since it's a union of finitely many countable sets.

Hence, $\bigcap\limits_{k=1}^{n} U_k \in \mathcal{T}$.

Exercise 13.4:

(a) If $\{\mathcal{T}_\alpha\}_{\alpha \in A}$ is a family of topologies on X, show that $\bigcap \mathcal{T}_\alpha$ is a topology on X. Is $\bigcup \mathcal{T}_\alpha$ a topology on X?

Let $\mathcal{T} = \bigcap\limits_{\alpha \in A} \mathcal{T}_\alpha$.

Since $\emptyset$ and X belong to all $\mathcal{T}_\alpha$, we know that $\emptyset, X \in \mathcal{T}$.

Next, suppose $\{U_\beta\}_{\beta \in B}$ is a collection of sets in $\mathcal{T}$. Since $\{U_\beta\}_{\beta \in B} \subseteq \mathcal{T}_\alpha$ for all α, we know that $\bigcup\limits_{\beta \in B} U_\beta \in \mathcal{T}_\alpha$ for all α. Hence, $\bigcup\limits_{\beta \in B} U_\beta \in \bigcap\limits_{\alpha \in A} \mathcal{T}_\alpha = \mathcal{T}$.

The same argument as used for arbitrary unions also shows that any finite intersection of sets in $\mathcal{T}$ is also in $\mathcal{T}$.

We've now shown that $\mathcal{T}$ is a topology. As for the other question asked, no we don't necessarily have that $\bigcup\limits_{\alpha\in A} \mathcal{T}_\alpha$ is a topology.

To see this, consider the set $X = \{a, b, c\}$ with the topologies $\mathcal{T}_1 = \{\emptyset, \{a\}, \{a, b\}, \{a, b, c\}\}$ and $\mathcal{T}_2 = \{\emptyset, \{c\}, \{b, c\}, \{a, b, c\}\}$. Then $\mathcal{T}_1 \cup \mathcal{T}_2$ is not a topology because $\{a\}, \{c\} \in \mathcal{T}_1 \cup \mathcal{T}_2$ but $\{a, c\} \notin \mathcal{T}_1 \cup \mathcal{T}_2$.

(b) Let $\{\mathcal{T}_\alpha\}_{\alpha\in A}$ be a family of topologies on X. Show that there is a unique smallest topology on X containing all the collections $\mathcal{T}_\alpha$, and a unique largest topology contained in all $\mathcal{T}_\alpha$.

Firstly, let $\{\mathcal{T}'_\beta\}_{\beta\in B}$ be the collection of all topologies on X which contain $\bigcup\limits_{\alpha\in A} \mathcal{T}_\alpha$.

We know that $\{\mathcal{T}'_\beta\}_{\beta\in B}$ is not empty because it must at least have $\mathcal{P}(X)$ as an element. Hence, we can apply part (a) of the problem to know that $\bigcap\limits_{\beta\in B} \mathcal{T}'_\beta$ is a topology on X.

Importantly, by virtue of being an intersection, that topology is smaller than all other topologies containing $\bigcup\limits_{\alpha\in A} \mathcal{T}_\alpha$. At the same time, we know it contains $\bigcup\limits_{\alpha\in A} \mathcal{T}_\alpha$.

So it is the unique smallest topology on X containing all the collections $\mathcal{T}_\alpha$.

The second part of this question is trivial from part (a). If a topology $\mathcal{T}''$ is contained in all $\mathcal{T}_\alpha$, then we know that $\mathcal{T}'' \subseteq \bigcap\limits_{\alpha\in A} \mathcal{T}_\alpha$. Clearly, the largest topology satisfying this is $\bigcap\limits_{\alpha\in A} \mathcal{T}_\alpha$.

Exercise 13.5: Show that if $\mathcal{A}$ is a basis for a topology on X, then the topology generated by $\mathcal{A}$ equals the intersection of all topologies on X that contain $\mathcal{A}$.

Let $\mathcal{T}$ be the topology generated by $\mathcal{A}$, and suppose $\mathcal{T}'$ is any topology containing $\mathcal{A}$. Then consider any $U \in \mathcal{T}$. By Lemma 13.1, we know that $U = \bigcup\limits_{\beta\in B} A_\beta$ where $\{A_\beta\}$ is some collection of sets in $\mathcal{A}$. Hence, U is a union of sets in $\mathcal{T}'$, meaning $U \in \mathcal{T}'$.

Since $\mathcal{T} \subseteq \mathcal{T}'$ for all $\mathcal{T}'$ containing $\mathcal{A}$, we thus know that $\mathcal{T}$ is the unique smallest topology containing $\mathcal{A}$. At the same time, by exercise 13.4.a, we know that the intersection $\mathcal{T}''$ of all topologies containing $\mathcal{A}$ is a topology. By virtue of being an intersection, we know it is smaller than all topologies containing $\mathcal{A}$, and that it contains $\mathcal{A}$. So, $\mathcal{T} \subseteq \mathcal{T}'' \subseteq \mathcal{T} \Longrightarrow \mathcal{T} = \mathcal{T}''$.

Prove the same if $\mathcal{A}$ is a subbasis.

Let $\mathcal{T}$ be the topology generated by $\mathcal{A}$ and suppose $\mathcal{T}'$ is any topology containing $\mathcal{A}$. Then consider any $U \in \mathcal{T}$. We know that $U = \bigcup_{\beta \in B} U_\beta$ where $\{U_\beta\}_{\beta \in B}$ is a collection of finite intersections of sets in $\mathcal{A}$.

Because each U_β must be in $\mathcal{T}'$, we thus know that $U \in \mathcal{T}'$. So $\mathcal{T} \subseteq \mathcal{T}'$.

The rest of the proof goes exactly the same as before.

This fact can be used as a shortcut for finding the unique smallest topology containing all topologies in a collection.

Exercise 13.1: Let X be a topological space and let A be a subset of X. Suppose that for each $x \in A$, there is an open set U such that $x \in U \subseteq A$. Then A is open in X.

For all $x \in A$, pick an open set U_x such that $x \in U_x \subseteq A$. Then $A = \bigcup_{x \in A} U_x$ is a union of open sets.

(A.O.C. usage!!)

9/27/2024

If X is simply-ordered, the standard topology for X (called the order topology) is defined as follows:

Given any $a, b \in X$ with $a < b$, we define the sets (a, b), $[a, b)$, $(a, b]$ and $[a, b]$ as you would expect. These are the open, closed, and half-open intervals on X.

Now let $\mathcal{B}$ be the collection of all sets of the form:

- Open intervals (a, b) in X.
- Intervals of the form $[a_0, b)$ where a_0 is the smallest element of X (if one exists).
- Intervals of the form $(a, b_0]$ where b_0 is the largest element of X (if one exists).

The collection $\mathcal{B}$ is a basis for a topology on X which is called the order topology.

It's fairly trivial to show that this is a basis. It's just that for the second condition of a basis, there are a bunch of cases that need to be mentioned.

Another way we can define the order topology is through rays. Given any $a \in X$, we define the sets $(a, +\infty)$, $(-\infty, a)$, $[a, +\infty)$ and $(-\infty, a]$ as you would expect.

Let $\mathcal{S}$ be the set of open rays: $(a, +\infty)$ and $(-\infty, a)$. This is a subbasis for the order topology on X.

To see this, firstly note that all open rays are open sets in the order topology of X. So, every set in the topology generated by $\mathcal{S}$ will be an open set in our original order topology. Hence if $\mathcal{T}$ is the order topology on X and $\mathcal{T}'$ is the topology generated by S, we know that $\mathcal{T}' \subseteq \mathcal{T}$.

At the same time, every interval in the previously defined basis of $\mathcal{T}$ is the intersection of two (or one if the interval contains the greatest or least element of X) rays. Hence, $\mathcal{T} \subseteq \mathcal{T}'$.

9/29/2024

Today I'm going to be studying from Michael Artin's textbook *Algebra, second edition*. My reasoning is that I need to learn more group theory in order to be ready for 188.

A law of compositon on a set S is map from $S \times S$ to S. Given the ordered pair $(a, b) \in S \times S$, we denote the element the pair is mapped to as either ab, $a \times b$, $a \circ b$, $a + b$, or etc.

Typically, $+$ is used if the composition is commutative. Meanwhile, the multiplicative symbols don't imply commutativity.

Proposition 2.1.4: Let an associative law of composition be given on S. Then we can uniquely define for all $n \in \mathbb{N}$ a product of n elements $a_1, \ldots, a_n$ of S, temporarily denoted by $[a_1 \cdots a_n]$, with the following properties:

(i) The product $[a_1]$ of one element is a_1.

(ii) The product $[a_1 a_2]$ of two elements is given by the law of composition.

(iii) For any integer i in the range $1 \leq i < n$, $[a_1 \cdots a_n] = [a_1 \cdots a_i][a_{i+1} \cdots a_n]$.

Proof:
We proceed by induction on n.

Let us define $[a_1 \cdots a_n] = [a_1 \cdots a_{n-1}][a_n]$ and suppose that the analogous definition of $[a_1 \cdots a_r]$ satisfies our properties for all $1 < r < n$. Then for any $1 \leq i < n - 1$, we have that:

$$\begin{aligned}
[a_1 \cdots a_n] &= [a_1 \cdots a_{n-1}][a_n] && \text{(by definition)} \\
&= ([a_1 \cdots a_i][a_{i+1} \cdots a_{n-1}])[a_n] && \text{(by inductive hypothesis)} \\
&= [a_1 \cdots a_i]([a_{i+1} \cdots a_{n-1}][a_n]) && \text{(by associativity)} \\
&= [a_1 \cdots a_i][a_{i+1} \cdots a_{n-1} a_n]
\end{aligned}$$

Based on the previous proposition, it's safe to just denote the product of $a_1, \ldots, a_n$ as $a_1 \cdots a_n$.

An identity for a law of composition is an element e of S satisfying that:

$ea = a$ and $ae = a$ for all $a \in S$.

We denote the identity of a law of composition as 0 or 1 (depending on whether we are using multiplicative or additive notation). We can only have one identity element.

Proof:
Suppose e and e' are both identity elements. Then $e = ee' = e'$.

An element a of S is invertible if there is another element $b \in S$ such that $ab = 1$ and $ba = 1$. We call b the inverse of a and denote b as $-a$ or a^{-1} depending on whether additive or multiplicative notation is being used.

Exercise 1.2:

- If an element a has both a left inverse l and a right inverse r, then $l = r$, a is invertible, and r is its inverse.

  Suppose $la = 1$ and $ar = 1$. Then we have that:
  $$r = 1r = (la)r = l(ar) = l1 = l$$

- If a is invertible, its inverse is unique.

  Suppose b and b' are both inverses of a. Then:
  $$b = 1b = (b'a)b = b'(ab) = b'1 = b'$$

- If a and b are invertible, then ab is invertible with $(ab)^{-1} = b^{-1}a^{-1}$.

  Proof:
  $abb^{-1}a^{-1} = a1a^{-1} = aa^{-1} = 1$ and $b^{-1}a^{-1}ab = b^{-1}1b = b^{-1}b = 1$

A group is a set G together with a law of composition such that:

1. The law of composition is associative.
2. G has an identity element.
3. Every element of G has an inverse.

An abelian group is a group whose law of composition is commutative.

The order of a group G is the number of elements it contains. We denote the order $|G|$. If $|G|$ is finite, we say G is a finite group. Otherwise, we say G is an infinite group.

The $n \times n$ general linear group is the group of all invertible $n \times n$ matrices. It's denoted GL_n. If we want to specify whether we are working with real or complex matrices, we write $GL_n(\mathbb{R})$ or $GL_n(\mathbb{C})$.

The symmetric group of a set is the set of permutations of the set (with the law of composition being function composition). We denote S_n the group of permutations of the indices $1, 2, \dots n$.

S_3 has order 6 (this is an easy fact from combinatorics). Now let $x = (1\ \ 2\ \ 3)$ and $y = (1\ \ 2)(3)$ (this is using cyclic notation). Then we can express all the elements of S_3 as products of x and y.

$$\begin{array}{lll} 1 = (1)(2)(3) & x = (1\ \ 2\ \ 3) & x^2 = (1\ \ 3\ \ 2) \\ y = (1\ \ 2)(3) & xy = (1\ \ 3)(2) & x^2y = (1)(2\ \ 3) \end{array}$$

Also note that $x^3 = 1$, $y^2 = 1$, and $yx = x^2y$.

Exercise 2.1: Make a multiplication table for the symmetric group S_3.

Just to clarify, to read this table, take the row element to be on the left side of the product and the column element to be on the right side of the product.

$\times$	1	x	x^2	y	xy	x^2y
1	1	x	x^2	y	xy	x^2y
x	x	x^2	1	xy	x^2y	y
x^2	x^2	1	x	x^2y	y	xy
y	y	x^2y	xy	1	x^2	x
xy	xy	y	x^2y	x	1	x^2
x^2y	x^2y	xy	y	x^2	x	1

12/14/2024

My goal for while I'm flying home and can't work on grading is to go prove the following extra exercise of homework 2 in my math 188 class.

Let $V = \{F(x) \in \mathbb{C}[[x]] : F(0) = 0\}$ and $W = \{F(x) \in \mathbb{C}[[x]] : G(0) = 1\}$.

(a) Given $F(x) \in V$, show that $\boldsymbol{E}(F(x)) = \sum_{n \geq 0} \frac{F(x)^n}{n!}$ is the *unique* formal power series $G(x) \in W$ such that $\mathrm{D}G(x) = \mathrm{D}F(x) \cdot G(x)$. This defines a function $\boldsymbol{E} : V \longrightarrow W$.

Note that we use the convention $F(x)^0 = 1$ even if $F(x) = 0$.

Firstly, note that we have $G(x) \coloneqq \boldsymbol{E}(F(x)) = \exp(F(x))$ where $\exp(x) \coloneqq \sum_{n\geq 0} \frac{x^n}{n!}$. Also, you can check $\exp(x)$ is it's own derivative. Thus by chain rule:

$$G(x) = \mathrm{D}F(x) \cdot \mathrm{D}(\exp)(F(x)) = \mathrm{D}F(x) \cdot \exp(F(x)) = \mathrm{D}F(x) \cdot G(x)$$

Next, suppose $H(x)$ is another formal power series in W satisfying that $\mathrm{D}H(x) = \mathrm{D}F(x) \cdot H(x)$. Note that since $H(0) \neq 0 \neq G(0)$, we can write that $\frac{\mathrm{D}H(x)}{H(x)} = \mathrm{D}F(x) = \frac{\mathrm{D}G(x)}{G(x)}$. Therefore, we get that:

$$(*) \quad \mathrm{D}H(x) \cdot G(x) = \mathrm{D}G(x) = H(x)$$

Let $H(x) = \sum_{n\geq 0} h_n x^n$ and $G(x) = \sum_{n\geq 0} g_n x^n$. Since we assumed that $H(x), G(x) \in W$, we know that $h_0 = g_0 = 1$. Then, proceeding by induction (assuming that $h_i = g_i$ for all $0 \leq i \leq n$), when we take the nth. coefficient of $(*)$ we get:

$$\begin{aligned}(n+1)h_{n+1} + \sum_{i=0}^{n-1}(i+1)h_{i+1}g_{n-i} \\ = \sum_{i=0}^{n}(i+1)h_{i+1}g_{n-i} = \sum_{i=0}^{n}(i+1)g_{i+1}h_{n-i} \\ = (n+1)g_{n+1} + \sum_{i=0}^{n-1}(i+1)g_{i+1}h_{n-i}\end{aligned}$$

But by induction we have $\sum_{i=0}^{n-1}(i+1)g_{i+1}h_{n-i} = \sum_{i=1}^{n-1}(i+1)h_{i+1}g_{n-i}$.

So subtracting out the sum from $i = 0$ to $n-1$ and then dividing by $n+1$ which is crucially nonzero, we then have that $h_{n+1} = g_{n+1}$.

(b) Given $G(x) \in W$, show that there is a *unique* formal power series $F(x) \in V$ such that $\mathrm{D}F(x) = \frac{\mathrm{D}G(x)}{G(x)}$. This let's us define the function $\boldsymbol{L} : W \longrightarrow V$ by $\boldsymbol{L}(G(x)) = F(x)$.

Since $G(0) = 1$, we know that $G(x)$ is invertible. So there is a unique formal power series $A(x) = \sum_{n\geq 0} a_n x^n$ such that $A(x) = \frac{\mathrm{D}G(x)}{G(x)}$.

Then if $F(x) = \sum_{n\geq 0} f_n x^n$ satisfies that $\mathrm{D}F(x) = \frac{\mathrm{D}G(x)}{G(x)}$, then we can solve that $f_n = \frac{a_{n-1}}{n}$ for all $n \geq 1$. This shows that f_n is uniquely determined for all $n \geq 1$. Also, since we are forcing $F(0) = 0$, we know that $f_0 = 0$. So $F(x)$ is a unique power series.

From this it's also hopefully clear to see how one can solve for $F(x)$ in order to show that $F(x)$ exists.

(c) Show that $\boldsymbol{E}$ and $\boldsymbol{L}$ are inverses of each other.

Firstly, we'll show $\boldsymbol{L}(\boldsymbol{E}(F(x))) = F(x)$ for all $F(x) \in V$.

Let $F(x) \in V$, $G(x) = \boldsymbol{E}(F(x))$, and $H(x) = \boldsymbol{L}(G(x))$. Then we have that:

$$\mathrm{D}H(x) = \frac{\mathrm{D}G(x)}{G(x)} = \frac{\mathrm{D}F(x) \cdot G(x)}{G(x)} = \mathrm{D}F(x)$$

This proves that $[x^n]H(x) = [x^n]F(x)$ for all $n \geq 1$. And since both $H(x), F(x) \in V$, we know that $H(0) = F(0)$. so $H(x) = F(x)$.

Secondly, we'll show $\boldsymbol{E}(\boldsymbol{L}(F(x))) = F(x)$ for all $F(x) \in W$.

Let $F(x) \in V$, $G(x) = \boldsymbol{L}(F(x))$, and $H(x) = \boldsymbol{E}(G(x))$. Then we have that:

$$\mathrm{D}H(x) = \mathrm{D}G(x) \cdot H(x) = \frac{\mathrm{D}F(x)}{F(x)} \cdot H(x)$$

Thus we know that $\mathrm{D}H(x) \cdot F(x) = \mathrm{D}F(x) \cdot H(x)$. Since both $H(x)$ and $F(x)$ are in W, we can employ identical logic as that of part (a) to show that $H(x) = F(x)$.

(d) Show that $\boldsymbol{E}(F_1(x) + F_2(x)) = \boldsymbol{E}(F_1(x))\boldsymbol{E}(F_2(x))$ for all $F_1(x), F_2(x) \in V$.

Note that:

$$\begin{aligned}
\boldsymbol{E}(F_1(x) + F_2(x)) &= \lim_{m\to\infty} \sum_{n=0}^{m} \frac{(F_1(x)+F_2(x))^n}{n!} \\
&= \lim_{m\to\infty} \sum_{n=0}^{m} \frac{1}{n!} \sum_{i=0}^{n} \frac{n!}{i!(n-i)!} F_1(x)^i F_2(x)^{n-i} \\
&= \lim_{m\to\infty} \sum_{n=0}^{m} \sum_{i=0}^{n} \frac{F_1(x)^i}{i!} \cdot \frac{F_2(x)^{n-i}}{(n-i)!} \\
&= \lim_{m\to\infty} \left((\sum_{n=0}^{m} \frac{F_1(x)^n}{n!})(\sum_{n=0}^{m} \frac{F_2(x)^n}{n!}) + R_m(x) \right)
\end{aligned}$$

In the above manipulations $R_m(x)$ is the formal power series negating all the terms which are in $(\sum_{n=0}^{m} \frac{F_1(x)^n}{n!})(\sum_{n=0}^{m} \frac{F_2(x)^n}{n!})$ but aren't in $\sum_{n=0}^{m} \sum_{i=0}^{n} \frac{F_1(x)^i}{i!} \cdot \frac{F_2(x)^{n-i}}{(n-i)!}$.

In other words, $R_m(x)$ contains all the terms of the form $\frac{1}{i!j!} F_1(x)^i F_2(x)^j$ where $i + j > m$. Importantly, because $F_1(0) = 0 = F_2(0)$, we know that $\mathrm{mdeg}\ R(x) > m$. So, $R_m(x) \to 0$ as $m \to \infty$.

In turn:

$$\lim_{m\to\infty} \left((\sum_{n=0}^{m} \frac{F_1(x)^n}{n!})(\sum_{n=0}^{m} \frac{F_2(x)^n}{n!}) + R_m(x) \right)$$

$$= \lim_{m\to\infty} \sum_{n=0}^{m} \frac{F_1(x)^n}{n!} \cdot \lim_{m\to\infty} \sum_{n=0}^{m} \frac{F_2(x)^n}{n!} + \lim_{m\to\infty} R_m(x) = \boldsymbol{E}(F_1(x))\boldsymbol{E}(F_2(x))$$

(e) Show that $\boldsymbol{L}(F_1(x)F_2(x)) = \boldsymbol{L}(F_1(x)) + \boldsymbol{L}(F_2(x))$ for all $F_1(x), F_2(x) \in W$.

By part (d), we know that:

- $\boldsymbol{E}(\boldsymbol{L}(F_1(x)) + \boldsymbol{L}(F_2(x))) = \boldsymbol{E}(\boldsymbol{L}(F_1(x)))\boldsymbol{E}(\boldsymbol{L}(F_2(x)))$

Meanwhile, by part (c) we know that:

- $\boldsymbol{E}(\boldsymbol{L}(F_1(x)F_2(x))) = F_1(x)F_2(x)$
- $\boldsymbol{E}(\boldsymbol{L}(F_1(x)))\boldsymbol{E}(\boldsymbol{L}(F_2(x))) = F_1(x)F_2(x)$

Hence, we've shown that $\boldsymbol{E}$ maps the left- and right-hand sides of the above claimed equation to the same formal power series. But from part (c) we know that $\boldsymbol{E}$ is an injective map. So we must have that the two sides of the equation are in fact equal.

(f) If m is a positive integer and $G(x) \in W$, show that $\boldsymbol{E}(\frac{\boldsymbol{L}(G(x))}{m})$ is an mth. root of $G(x)$.

Let $F(x) = \boldsymbol{E}(\frac{\boldsymbol{L}(G(x))}{m})$. This implies $m \cdot \boldsymbol{L}(F(x)) = \boldsymbol{L}(G(x))$. Then by part (e), we know that $\boldsymbol{L}(F(x)^m) = \boldsymbol{L}(G(x))$. And finally, plugging both sides into $\boldsymbol{E}$ we get $F(x)^m = G(x)$.

Also, $F(0) = 1$. So $F(x)$ is the unique mth. root of $G(x)$ with 1 as its constant coefficient.

Because of part (f), we can now extend our definition of the zth. power of a formal power series $G(x) \in W$ to any complex number z. Specifically, for any $G(x) \in W$ we define $G(x)^z := \boldsymbol{E}(z \cdot \boldsymbol{L}(G(x)))$. Then, we've shown in part (f) that this definition agrees with our more restricted definition from math 188 on at least the positive rational numbers.

Two important identities (where $G(x) \in W$ and $z, w \in \mathbb{C}$):

- $G(x)^z G(x)^w = G(x)^{z+w}$?

  Proof:
  $$\begin{aligned} G(x)^z G(x)^w &= \boldsymbol{E}(z \cdot \boldsymbol{L}(G(x)))\boldsymbol{E}(w \cdot \boldsymbol{L}(G(x))) \\ &= \boldsymbol{E}(z \cdot \boldsymbol{L}(G(x)) + w \cdot \boldsymbol{L}(G(x))) = \boldsymbol{E}((z+w) \cdot \boldsymbol{L}(G(x))) \\ &= G(x)^{z+w} \end{aligned}$$

- $(G(x)^z)^w = G(x)^{zw}$

  Proof:
  $$\begin{aligned} (G(x)^z)^w &= (\boldsymbol{E}(z \cdot \boldsymbol{L}(G(x))))^w \\ &= \boldsymbol{E}(w \cdot \boldsymbol{L}(\boldsymbol{E}(z \cdot \boldsymbol{L}(G(x))))) = \boldsymbol{E}(w \cdot (z \cdot \boldsymbol{L}(G(x)))) \\ &= \boldsymbol{E}(wz \cdot \boldsymbol{L}(G(x))) = G(x)^{zw} \end{aligned}$$

Using those identities, here are some corollaries:

- $G(x)^{-a}$ gives the multiplicative inverse of $G(x)^a$

  (this proves that this definition of exponentiation agrees with our more restricted definition from math 188 on all rational numbers).

  Proof:
  Firstly note that:
  $$G(x)^{-a}G(x)^a = G(x)^{-a+a} = G(x)^0 = \boldsymbol{E}(0 \cdot \boldsymbol{L}(G(x))) = \boldsymbol{E}(0)$$

  Seconly, note that $\boldsymbol{E}(0) = \sum_{n\geq 0} \frac{0^n}{n!} = \frac{0^0}{0!} = 1$ (as a reminder we are using the convention that $0^0 = 1$). Therefore, $G(x)^{-a}G(x)^a = 1$, meaning $G(x)^{-a}$ is the multiplicative inverse of $G(x)^a$.

- V is a subgroup of $\mathbb{C}[[x]]$ under addition, W is a group under multiplication, and $\boldsymbol{E}$ is a group isomorphism between $(V, +)$ and $(W, \cdot)$.

  Proof:
  One can easily see without our prior reasoning that $(V, +)$ is subgroup of $\mathbb{C}[[x]]$ under addition (with 0 as it's identity).

  Meanwhile, by our previous corollary we can see that all $G(x) \in W$ have a multiplicative inverse in W. Specifically since $G(x) = G(x)^1$, we know by the previous corollary that $G(x)$ has the inverse $G(x)^{-1}$ inside W. Combining that with the fact that $1 \in W$ and $G(x)H(x) \in W$ when $G(x), H(x) \in W$, we know now that W is a group under multiplcation.

  Finally, we know $\boldsymbol{E}$ is a group isomorphism because of part (c) of this exercise as well as the fact that $\boldsymbol{E}(0) = 1$.

- Power Rule: Given $G(x) \in W$, if $H(x) = G(x)^z$, then
  $$\mathrm{D}H(x) = z\mathrm{D}G(x)G(x)^{z-1}.$$

  Proof:
  $$\begin{aligned}\mathrm{D}H(x) &= \mathrm{D}(\boldsymbol{E}(z \cdot \boldsymbol{L}(G(x))))(x) \\ &= \mathrm{D}(z \cdot \boldsymbol{L}(G(x)))(x) \cdot \boldsymbol{E}(z \cdot \boldsymbol{L}(G(x))) \\ &= z \cdot \mathrm{D}(\boldsymbol{L}(G(x)))(x) \cdot G(x)^z = z\frac{\mathrm{D}G(x)}{G(x)}G(x)^z \\ &\qquad\qquad = z\mathrm{D}G(x)G(x)^{-1}G(x)^z \\ &\qquad\qquad = z\mathrm{D}G(x)G(x)^{z-1}\end{aligned}$$

- Binomial Theorem: Given $z \in \mathbb{C}$, we have that:
  $$(1+x)^z = \sum_{n\geq 0} \binom{z}{n} x^n$$ where $\binom{z}{n} = \frac{z(z-1)\cdots(z-n+1)}{n!}$ when $n > 0$ and 1 when $n = 0$.

  Proof:
  Note that $[x^n](1+x)^z = \frac{1}{n!}\mathrm{D}^n((1+x)^z)(0)$. Also, by induction using the power rule we can say for $n > 0$ that:

$$\begin{aligned}\mathrm{D}^n((1+x)^z)(x) &= z\mathrm{D}^{n-1}((1+x)^z)(x) \\ &= z(z-1)\mathrm{D}^{n-2}((1+x)^z) \\ &= \cdots = z(z-1)\cdots(z-n+1)(1+x)\end{aligned}$$

Therefore $\mathrm{D}^n((1+x)^z)(0) = z(z-1)\cdots(z-n+1)(1+0)$ and we thus have that for $n > 0$.

$$[x^n](1+x)^z = \tfrac{1}{n!}z(z-1)\cdots(z-n+1) = \tbinom{z}{n}.$$

Meanwhile, if $n = 0$, then $[x^0](1+x)^z = 1 = \binom{z}{n}$ (because $(1+x)^z \in W$).

Before going on to parts (g) and (h), here are two more identities (where $F(x) \in V$, $G(x) \in W$, and $z \in \mathbb{C}$):

- $(\boldsymbol{E}(F(x)))^z = \boldsymbol{E}(z \cdot \boldsymbol{L}(\boldsymbol{E}(F(x)))) = \boldsymbol{E}(zF(x))$
- $\boldsymbol{L}(G(x)^z) = \boldsymbol{L}(\boldsymbol{E}(z \cdot \boldsymbol{L}(G(x)))) = z\boldsymbol{L}(G(x))$

(g) Show that if $\sum_{i\geq 0} F_i(x)$ converges to $F(x)$, then $\prod_{i\geq 0} \boldsymbol{E}(F_i(x))$ converges to $\boldsymbol{E}(F(x))$.

We start by proving the following lemma: If $A(x) \in \mathbb{C}[[x]]$ and $(B_i(x))_{i\in\mathbb{N}}$ is a sequence in $\mathbb{C}[[x]]$ converging to $B(x)$ as $i \to \infty$ and satisfying that $B_i(0) = 0$ for all i, then $A(B_i(x)) \to A(B(x))$ as $i \to \infty$.

Proof:
For notation, we'll denote:

$$A(x) = \textstyle\sum_{n\geq 0} a_n x^n, \quad B_i(x) = \sum_{n\geq 0} b_n^{(i)} x^n, \text{ and } B(x) = \sum_{n\geq 0} b_n x^n$$

To start, note that for all integers $m \geq 0$, we have that $B_i(x)^m \to B(x)^m$. Also, since $B_i(0) = 0$ for all i, we know that $\operatorname{mdeg} B_i(x)^m \geq m$ for all integers i and m, and also that $\operatorname{mdeg} B(x)^m \geq m$ for all integers m. Thus, fixing $n \geq 0$ we can say that:

$$[x^n]A(B_i(x)) = [x^n]\sum_{m=0}^{n} a_m B_i(x)^m \text{ and } [x^n]A(B(x)) = [x^n]\sum_{m=0}^{n} a_m B(x)^m$$

Next, let I_m be large enough that $[x^n]B_i(x)^m = [x^n]B(x)^m$ for all $i \geq I_m$. Then set $I = \max(I_0, I_1, \ldots, I_m)$ and note that for all $i \geq I$, we have:

$$\begin{aligned}[x^n]\sum_{m=0}^{n} a_m B_i(x)^m &= \sum_{m=0}^{n} a_m [x^n](B_i(x)^m) \\ &= \sum_{m=0}^{n} a_m [x^n](B(x)^m) = [x^n]\sum_{m=0}^{n} a_m B(x)^m\end{aligned}$$

So for all $i \geq I$, we have that $[x^n]A(B_i(x)) = [x^n]A(B(x))$. This proves $A(B_i(x)) \to A(B(x))$.

I should have proved this in my math 188 notes when I was showing that $((A+B)\circ C)(x) = (A\circ C)(x) + (B\circ C)(x)$ and $(AB\circ C)(x) = (A\circ C)(x)(B\circ C)(x)$. But in my defense the professor didn't mention any of these three facts in his notes.

As a reminder: $\boldsymbol{E}(F(x)) = \exp(F(x))$ where $\exp = \sum_{n\geq 0} \frac{1}{n!}x^n$. Also, by our previous lemma, we know that:

$$\exp(F(x)) = \exp(\lim_{n\to\infty}\sum_{i=0}^{n} F_i(x)) = \lim_{n\to\infty}\exp(\sum_{i=0}^{n} F_i(x))$$

But then $\exp(\sum_{i=0}^{n} F_i(x)) = \boldsymbol{E}(\sum_{i=0}^{n} F_i(x)) = \prod_{i=0}^{n} \boldsymbol{E}(F_i(x))$.

So, we have shown that $\boldsymbol{E}(F(x)) = \lim_{n\to\infty}\prod_{i=0}^{n} \boldsymbol{E}(F_i(x)) = \prod_{i\geq 0} \boldsymbol{E}(F_i(x))$

Side note: The lemma we proved in this part also tells us that if $(B_i(x))_{i\in\mathbb{N}}$ is a sequence in V converging to $B(x)$, then $\boldsymbol{E}(B_i(x)) \to \boldsymbol{E}(B(x))$ as $i \to \infty$. In other words, $\boldsymbol{E}$ is a continuous map.

(If $\rho(A(x), B(x)) = \frac{1}{\text{mdeg }(A-B)(x)}$, then $(\mathbb{C}[[x]], \rho)$ is a metric space in which formal power series convergence is equivalent to convergence in this metric space.)

(h) Show that if $\prod_{i\geq 0} G_i(x)$ converges to $G(x)$, then $\sum_{i\geq 0} \boldsymbol{L}(G_i(x))$ converges to $\boldsymbol{L}(G(x))$.

Unfortunately, unlike with $\boldsymbol{E}$ we do not (currently) have a formal power series $A(x)$ for which we can generally say $\boldsymbol{L}(B(x)) = A(B(x))$. Thus, we can't move the limit from inside $\boldsymbol{L}$ to outside $\boldsymbol{L}$ as easily as we did in part (g) for $\boldsymbol{E}$.

However, consider that $\lim_{n\to\infty} \boldsymbol{E}(\sum_{i=0}^{n} \boldsymbol{L}(G_i(x)))$ exists. Specifically:

$$\begin{aligned}\lim_{n\to\infty} \boldsymbol{E}(\sum_{i=0}^{n} \boldsymbol{L}(G_i(x))) &= \lim_{n\to\infty} \boldsymbol{E}(\sum_{i=0}^{n} \boldsymbol{L}(G_i(x))) \\ &= \lim_{n\to\infty}\prod_{i=0}^{n} \boldsymbol{E}(\boldsymbol{L}(G_i(x))) = \lim_{n\to\infty}\prod_{i=0}^{n} G_i(x) = G(x)\end{aligned}$$

Thus, if we can show $\sum_{i\geq 0} \boldsymbol{L}(G_i(x))$ converges, then we can use the lemma from part (g) to see that:

$$\boldsymbol{E}(\lim_{n\to\infty}\sum_{i=0}^{n} \boldsymbol{L}(G_i(x))) = \lim_{n\to\infty} \boldsymbol{E}(\sum_{i=0}^{n} \boldsymbol{L}(G_i(x))) = G(x),$$

and then by applying $\boldsymbol{L}$ to the left and right sides of this equation, we will get our desired result.

We now show that $\sum_{i\geq 0} \boldsymbol{L}(G_i(x))$ converges. Firstly, note that after fixing $n \geq 1$, we have that:

$$[x^n](\boldsymbol{L}(G_i(x)))(x) = \tfrac{1}{n}[x^{n-1}]\tfrac{\mathrm{D}G_i(x)}{G_i(x)}.$$

Secondly, note that $\prod_{i\geq 0} G_i(x)$ converging to $G(x)$ and $G_i(0) = 1$ for all i implies that $\operatorname{mdeg}(G_i(x) - 1) \to \infty$ as $i \to \infty$ and thus $\operatorname{mdeg} \mathrm{D}G_i(x) = \operatorname{mdeg}(G_i(x) - 1) - 1 \to \infty$ as $i \to \infty$.

Hence, there exists $I_n \geq 0$ such that $i \geq I_n$ implies that $[x^0]\mathrm{D}G_i(x)$, $[x^1]\mathrm{D}G_i(x), \ldots, [x^{n-1}]\mathrm{D}G_i(x)$ are all 0. In turn, we have for all $i \geq I_n$ that $[x^{n-1}](\mathrm{D}G_i(x) \cdot \frac{1}{G_i(x)}) = 0$ and $[x^n](\boldsymbol{L}(G_i(x)))(x) = \frac{1}{n} \cdot 0 = 0$.

Since we also have by the definition of $\boldsymbol{L}$ that $(\boldsymbol{L}(G_i(x)))(0) = 0$ for all i, we can thus conclude that: $\lim\limits_{i\to\infty} \operatorname{mdeg}(\boldsymbol{L}(G_i(x))) = 0$.

This proves that $\sum_{i\geq 0} \boldsymbol{L}(G_i(x))$ converges.

To finish off, here is an identity: If $z, \alpha \in \mathbb{C}$ and $G(x) = (1 - \alpha x)^z$, then $\boldsymbol{L}(G(x)) = \sum\limits_{n\geq 1} \frac{-z}{n}\alpha^n x^n$.

Proof:
We already showed that $\boldsymbol{L}((1-\alpha x)^z) = z \cdot \boldsymbol{L}(1 - \alpha x)$. Also:

$$\mathrm{D}(\boldsymbol{L}(1-\alpha x))(x) = \frac{\mathrm{D}(1-\alpha x)}{1-\alpha x} = -\alpha \sum_{n\geq 0} \alpha^n x^n = \sum_{n\geq 0} -\alpha^{n+1}x^n$$

Thus, we have that $z \cdot \boldsymbol{L}(1-\alpha x) = z \cdot \sum\limits_{n\geq 1} \frac{-1}{n}\alpha^{(n+1)-1}x^n = \sum\limits_{n\geq 1} \frac{-z}{n}\alpha^n x^n$.

Some thoughts on this:

- $\boldsymbol{L}(\frac{1}{1-x}) = \sum_{n\geq 1} \frac{+1}{n}(1)^n x^n = \sum_{n\geq 1} \frac{1}{n}x^n$ which is importantly the same as what we found in class.
- Given that $G(x)$ is a polynomial with constant coefficient 1, we can combine this identity with part (e) to calculate $\boldsymbol{L}(G(x))$.

  Specifically, let $G(x) = (1 - \frac{1}{\gamma_1}x)\cdots(1 - \frac{1}{\gamma_k}x)$ where $\gamma_1, \ldots, \gamma_k$ are the roots of $G(x)$. Since $G(0) = 1$ by assumption, we know that $\gamma_j \neq 0$ for all j. Then we have that:

$$\begin{aligned}\boldsymbol{L}(G(x)) &= \boldsymbol{L}((1 - \tfrac{1}{\gamma_1}x)\cdots(1 - \tfrac{1}{\gamma_k}x)) \\ &= \boldsymbol{L}(1 - \tfrac{1}{\gamma_1}x) + \ldots + \boldsymbol{L}(1 - \tfrac{1}{\gamma_k}x) \\ &= \sum_{n\geq 1} \tfrac{-1}{n}\gamma_1^{-n}x^n + \ldots + \sum_{n\geq 1} \tfrac{-1}{n}\gamma_k^{-n}x^n = \sum_{n\geq 1} \tfrac{-1}{n}\left(\sum_{j=1}^{k} \gamma_j^{-n}\right) x^n\end{aligned}$$

And now I'm out of ideas of what else to do with this homework problem.

12/19/2024

For the next while I want to work through some of the exercises in Folland's *Real Analysis* about the Cantor set and Cantor function. Assume for this section that $\mathbb{R}$ is equipped with the standard metric ρ and that we are using the complete Lebesgue measure space $(\mathbb{R}, \mathcal{L}, m)$.

If I is a bounded interval and $\alpha \in (0,1)$, then call the open interval with the same midpoint as I and length equal to α times the length of I the "open middle αth" of I. If $(\alpha_j)_{j\in\mathbb{N}}$ is a sequence of numbers in $(0,1)$, then we can define a decreasing sequence $(K_j)_{j\in\mathbb{N}}$ of closed sets by setting $K_0 = [0,1]$ and obtaining K_j by removing the open middle α_jth from the intervals that make up K_{j-1}. Then $K = \bigcap_{j\in\mathbb{N}} K_j$ is called a generalized Cantor set.

The ordinary Cantor set C is obtained by setting all α_j equal to $1/3$.

Exercise 2.27: Let $K = \bigcap K_j$ be a generalized Cantor set created using the sequence $(\alpha_j)_{j\in\mathbb{N}}$ in $(0,1)$. Prove that K is compact, perfect (i.e. closed and has no isolated points), nowhere dense (i.e. not dense in any nonempty open set), and totally disconnected (i.e. the only connected subsets of K are single points).

- K is closed because it is an intersection of closed sets. Also, K is a bounded set in $\mathbb{R}$ because it is a subset of $[0,1]$. Thus, K is compact.

- Let $x \in K$. Then for any $\varepsilon > 0$, pick $J \in \mathbb{N}$ with $2^{-J} < \varepsilon$. Note that all intervals of K_j have a length at most 2^{-j}. After all, when going from K_{j-1} to K_j, we split all the intervals of K_{j-1} in half and then remove an additional amount of length determined by α_j. So, let I be the interval of K_J containing x. Then both endpoints of I are in K and also in $B(\varepsilon, x)$. And, at least one of those endpoints is not x. So x is a limit point of K. Since K is also closed, we have that K is perfect.

- Let $x, y \in K$ and without loss of generality assume $x < y$. Then we know there must exist some integer $J \in \mathbb{N}$ such that x and y are in different intervals of K_J. Afterall, as previously mentioned, points in the same interval of K_j are within 2^{-j} distance of each other. So if no such J exists, then $\rho(x,y) < 2^{-j}$ for all j, meaning $x = y$.

  We can specifically choose J to be the least integer such that x and y are in two different intervals of K_J. Then both x and y are in the same interval I of K_{J-1}, but the midpoint of that interval z is not in $K_J \subseteq K$ and $x < z < y$. By a theorem in 140A, this proves that K is totally disconnected since for all $x, y \in K$, $[x,y] \not\subseteq K$ unless $x = y$.

- Since K is perfect, we know that K is only dense on subsets of K. However, since all open sets in $\mathbb{R}$ are countable unions of open intervals and K contains no nonempty open intervals since K is totally disconnected, we know that K has no nonempty open subsets.

Trying to explicitely write out the bijection between $[0,1]$ and a generalized Cantor set $K = \bigcap K_j$ would be really time consuming and awkward. So I'm going to be more handwavey

We can define an injection from $\{0,1\}^\omega$ to K as follows:

Given $\boldsymbol{x} \in \{0,1\}^\omega$, we can define a convergent subsequence in K. Specifically, set $a_0 = 0$. Then recursively for $j > 0$, we know a_{j-1} falls into some interval I of K_{j-1}. Futhermore, we know that I gets split into two disjoint intervals I_0 and I_1 when going from K_{j-1} to K_j (take I_0 to be the lower interval). Then, let a_j be the left bound on I_n where n is the value at the jth index of $\boldsymbol{x}$.

Since the endpoints of the intervals in each K_j are all in the final intersection, we know that $(a_j)_{j\in\mathbb{N}}$ is a sequence contained in K. Also, $\rho(a_j, a_{j+1}) < 2^{-j}$ for all $j \geq 0$. From that you can easily work out that $(a_j)_{j\in\mathbb{N}}$ is Cauchy. Thus, since K is closed, we know that $(a_j)_{j\in\mathbb{N}}$ converges to some number $y \in K$.

The mapping $\boldsymbol{x} \mapsto y$ is injective because $\boldsymbol{x}$ uniquely determines which interval of K_j that y is in for all j (specifically the same interval as a_j for each j). If $\boldsymbol{x}'$ is another sequence of 0s and 1s mapped to y', and $\boldsymbol{x}$ and $\boldsymbol{x}'$ differ at position J, then y and y' will be in two different intervals of K_J. Since those intervals are disjoint, we know that $y \neq y'$.

It is possible to show that our above injection is also surjective. However, it's quicker to just say $\mathfrak{c} = \mathrm{card}(\{0,1\}^\omega) \leq \mathrm{card}(K) \leq \mathrm{card}(\mathbb{R}) = \mathfrak{c}$. Thus generalized Cantor sets have the cardinality of the continuum.

Finally, note that given a generalized Cantor set $K = \bigcap K_j$, because $m(K_1) < 1$ and $(K_j)_{j\in\mathbb{N}}$ is a decreasing sequence of sets, we know that:

$$m(K) = \lim_{j\to\infty} m(K_j) = \lim_{j\to\infty} 2^j \prod_{i=1}^{j} \tfrac{(1-\alpha_i)}{2} = \prod_{j=1}^{\infty}(1-a_j)$$

Exercise 2.32:

(a) Suppose $(a_j)_{j\in\mathbb{N}}$ is a sequence in $(0,1)$. $\prod_{j=1}^\infty (1-a_j) > 0$ if and only if $\sum_{j=1}^\infty a_j < \infty$.

To start, note that for all $x > 0$, we have that $0 \leq x \leq -\log(1-x)$. After all, $x + \log(1-x)$ equals 0 at $x = 0$. Also, it's derivative: $1 - \frac{1}{1-x}$, is negative for all $x > 0$. This tells us that $x + \log(1-x)$ is strictly decreasing as x increases, meaning that the difference of x and $-\log(1-x)$ is less than 0 for all $x > 0$.

This lets us conclude that if $\sum_{j=1}^{\infty} -\log(1-a_j)$ converges, then by comparison test we must also have that $\sum_{j=1}^{\infty} a_j$ converges.

Meanwhile, for all $x \in [0, 1/2)$ we have that $0 \leq -\log(1-x) \leq 2x$. To see this, note that $2x + \log(1-x)$ also equals 0 at $x = 0$. But it's derivative: $2 - \frac{1}{1-x}$, is positive for $x < 1/2$. This tells us that $2x + \log(1-x)$ is strictly increasing for $x \in (0, 1/2)$. So, the difference of $2x$ and $-\log(1-x)$ is greater than 0 for all $x \in (0, 1/2)$.

Importantly, if $\sum_{j=1}^{\infty} a_j$ converges, then we know that all a_j after a certain index J will be in the interval $(0, 1/2)$. Then, since the sum of the $-\log(1-a_j)$ for $j \leq J$ will be finite and since we can use comparison test on the remaining terms, we know that $\sum_{j=1}^{\infty} -\log(1-a_j)$ also converges.

In other words, we've shown that:

$$\sum_{j=1}^{\infty} a_j < \infty \text{ if and only if } \sum_{j=1}^{\infty} -\log(1-a_j) < \infty.$$

Next, note that $\sum_{j=1}^{\infty} -\log(1-a_j)$ converges if and only $\sum_{j=1}^{\infty} \log(1-a_j)$ converges.

Finally, consider that $\prod_{j=1}^{\infty}(1-a_j) > 0$ if and only if $\sum_{j=1}^{\infty} \log(1-a_j) > -\infty$.

If $\prod_{j=1}^{\infty}(1-a_j) = \alpha > 0$, then we know that $\log(\alpha)$ is a finite negative value. And because $\log$ is a continuous function, we know:

$$\begin{aligned}\log(\alpha) &= \log(\lim_{N\to\infty} \textstyle\prod_{j=1}^{N}(1-a_j)) \\ &= \lim_{N\to\infty} \log(\textstyle\prod_{j=1}^{N}(1-a_j)) = \displaystyle\lim_{N\to\infty} \sum_{j=1}^{N} \log(1-a_j) = \sum_{j=1}^{\infty} \log(1-a_j)\end{aligned}$$

Meanwhile, if $\prod_{j=1}^{\infty}(1-a_j) = \lim_{N\to\infty} \prod_{j=1}^{N}(1-a_j) = 0$, then we know that:

$$\sum_{j=1}^{\infty} \log(1-a_j) = \lim_{N\to\infty} \sum_{j=1}^{N} \log(1-a_j) = \lim_{N\to\infty} \log(\prod_{j=1}^{N}(1-a_j)) = -\infty$$

Note that we always have that $\prod_{j=1}^{\infty}(1-a_j) \in [0, 1)$.

(b) Given $\beta \in (0, 1)$, there exists a sequence $(a_j)_{j\in\mathbb{N}}$ in $(0, 1)$ such that $\prod_{j=1}^{\infty}(1-a_j) = \beta$.

Let $c_j = \frac{1}{2^j}(1-\beta) + \beta$ for all $j \in \mathbb{N}$. That way $(c_j)_{j\in\mathbb{N}}$ is a strictly decreasing sequence in $(0, 1)$ converging to β. Then, we want to make $\prod_{i=1}^{j}(1-a_j) = c_j$ for all j. To do this, set $a_j = 1 - \frac{c_j}{c_{j-1}}$ for all j. Because $0 < c_j < c_{j-1}$, we know that $\frac{c_j}{c_{j-1}} \in (0, 1)$. And thus, $a_j \in (0, 1)$ for all j as well and $\prod_{j=1}^{\infty}(1-a_j) = \beta$.

Letting $C = \bigcap C_j$ be the standard Cantor set (i.e. where all $\alpha_j = 1/3$), we now define the Cantor function:

> Note that if $x \in C$, then there exists a unique sequence $(a_j)_{j\in\mathbb{N}}$ with $x = \sum_{j=1}^{\infty} a_j \frac{1}{3^j}$ and all a_j equal to either 0 or 2. (This is because each choice of a_j as either 0 or 2 corresponds to which subinterval of C_j that x is in.) Let $f(x) = \sum_{j=1}^{\infty} b_j 2^{-j}$ where $b_j = \frac{a_j}{2}$. Note that $f(x)$ is the binary expansion of a number in $[0, 1]$.

Observe that for all $y \in [0, 1]$ there exists $x \in C$ with $f(x) = y$. Also, for $x_1, x_2 \in C$ with $x_1 < x_2$, we have that $f(x_1) \leq f(x_2)$ with equality if and only if x_1 and x_2 are the end points of a removed interval (thus making x_1 and x_2 correspond to the binary expansions $0.b_1b_2\ldots0\overline{1}$ and $0.b_1b_2\ldots1\overline{0}$). This allows us to continuously extend f to all $[0, 1]$ by making f constant on all the intervals between points of C.

Exercise 2.9: Let $f : [0, 1] \longrightarrow [0, 1]$ be the Cantor function, and let $g(x) = f(x) + x$.

(a) g is a bijection from $[0, 1]$ to $[0, 2]$, and $h = g^{-1}$ is continuous from $[0, 2]$ to $[0, 1]$.

> To start, note that g is easily checked to be a strictly increasing function. This proves both that g is injective and that g has a range of $[0, 2]$ since $g(0) = 0$ and $g(1) = 2$. Also, note that g is continuous since both $f(x)$ and x are continuous. Thus, by applying I.V.T, we can say that g is surjective from $[0, 1]$ to $[0, 2]$. This proves that g is a bijection.
>
> The proof that $h = g^{-1}$ is continuous works for any strictly increasing continuous function.
>
> > Given $y \in [0, 2]$ there exists $x \in [0, 1]$ such that $g(x) = y$.
> >
> > Now let $\varepsilon > 0$ and set $\alpha = \max(0, x - \varepsilon)$ and $\beta = \min(1, x + \varepsilon)$. Then for all $y' \in (g(\alpha), g(\beta))$, we have that $h(y') \in B(\varepsilon, x)$. So we can set $\delta = \min(|g(\alpha) - y|, |g(\beta) - y|)$. This fulfills the definition of continuity.

(b) $m(g(C)) = 1$ where $C = \bigcap C_j$ is the Cantor set.

> Because g^{-1} is continuous, we know that $g(C)$ and $g(C_j)$ are measurable for all j since C and each C_j are Borel sets. Also $(g(C_j))_{j\in\mathbb{N}}$ is a decreasing sequence of sets with $m(g(C_1)) \leq 2$ and $\bigcap_{j\in\mathbb{N}} g(C_j) = g(C)$. Thus $m(g(C)) = \lim\limits_{j\to\infty} m(g(C_j))$.
>
> Next note that C_j has 2^j many intervals, each with width 3^{-j}. Also, if $[\alpha, \alpha + 3^{-j}]$ is one of those intervals, then:
>
> $$\begin{aligned} g(\alpha + 3^{-j}) - g(\alpha) &= f(\alpha + 3^{-j}) + \alpha + 3^{-j} - f(\alpha) - 3^{-j} \\ &= \textstyle\sum_{i=1}^{j}(b_i 2^{-i}) + 2^{-j} + \alpha + 3^{-j} - \sum_{i=1}^{j}(b_i 2^{-i}) - \alpha \\ &= 2^{-j} + 3^{-j} \end{aligned}$$

Thus $m(g(C_j)) = 2^j(2^{-j} + 3^{-j}) = 1 + (\frac{2}{3})^j$. Taking $j \to \infty$ we get the desired result.

To do the next parts of that exercise, we first need to do a different exercise.

Exercise 1.29:

(a) Suppose $E \subseteq V$ where V is a Vitali set (see the tangent on page 22) and $E \in \mathcal{L}$. Prove that $m(E) = 0$.

For all $r \in \mathbb{Q} \cap [-1, 1]$, define $E_r = \{v + r : v \in E\}$. By translation invariance, we know that E_r is measurable with $m(E_r) = m(E)$ for all r. Also each E_r is disjoint and $\bigcup_{r \in \mathbb{Q} \cap [-1,1]} E_r \subseteq [-1, 2]$. It follows that $\bigcup_{r \in \mathbb{Q} \cap [-1,1]} E_r$ is measurable and:

$$3 \geq m(\bigcup_{r \in \mathbb{Q} \cap [-1,1]} E_r) = \sum_{r \in \mathbb{Q} \cap [-1,1]} m(E_r) = \sum_{r \in \mathbb{Q} \cap [-1,1]} m(E)$$

The only way this is possible is if $m(E) = 0$.

(b) If $m(E) > 0$, then there exists a nonmeasurable set $N \subseteq E$.

Sidenote: the converse of this statement is trivially true because $(\mathbb{R}, \mathcal{L}, m)$ is complete.

To start, it suffices to show this for $E \subseteq [0, 1]$. After all, we can use the translation invariance of the Lebesgue measure to move N from $[0, 1]$ to where ever E is (we can't have that $N + r$ is measurable because that would imply $(N + r) - r = N$ is measurable).

Now, let V be a Vitali set and $V_r = \{v + r : v \in V\}$ for all $r \in [-1, 1] \cap \mathbb{Q}$. If $E \cap V_r$ is not measurable for some r, then we are done. So, suppose $E \cap V_r$ is measurable for all r. Then note $\bigcup_{r \in [-1,1] \cap \mathbb{Q}} (E \cap V_r) = E \cap \bigcup_{r \in [-1,1] \cap \mathbb{Q}} (V_r)$. Since $[0, 1]$ is a subset of $\bigcup_{r \in [-1,1] \cap \mathbb{Q}} (V_r)$ and $E \subseteq [0, 1]$, we thus know that $E \cap \bigcup_{r \in [-1,1] \cap \mathbb{Q}} (V_r) = E$. Additionally, since each $E \cap V_r$ is disjoint, we know that:

$$m(E) = \sum_{r \in [-1,1] \cap \mathbb{Q}} m(E \cap V_r)$$

Now hopefully it's clear how part (a) of this exercise extends to each nonmeasurable set V_r. Thus, since we assumed each $E \cap V_r$ is a measurable set, we know that $m(E \cap V_r) = 0$. It follows that $m(E) = 0$, a contradiction of our problem.

Now we return to exercise 2.9.

(c) $g(C)$ contains a Lebesgue nonmeasurable set A by exercise 1.29. Let $B = g^{-1}(A)$. Then B is Lebesgue measurable but not Borel.

Since C is a measurable null set in the complete measure space $(\mathbb{R}, \mathcal{L}, m)$, we have that all subsets of C including B must be measurable. So $B \in \mathcal{L}$.

Side note: since $\mathrm{card}(C) = \mathrm{card}(\mathbb{R})$, we know that:

$$\mathrm{card}(\mathcal{P}(\mathbb{R})) = \mathrm{card}(\mathcal{P}(C)) \leq \mathrm{card}(\mathcal{L}) \leq \mathrm{card}(\mathcal{P}(\mathbb{R})).$$

However, because g^{-1} is continuous, we know that g^{-1} is a Borel measurable function. Hence, if B was borel, then we would have to have that $g(B) = A$ is also Borel, thus contradicting that A is not measurable. So, we know B is measurable but not Borel.

(d) There exists a Lebesgue measurable function F and continuous function G on $\mathbb{R}$ such that $F \circ G$ is not Lebesgue measurable.

Define G by continuously extending $g^{-1}(x)$ to all $\mathbb{R}$ (One way to do this would be to set $G(x) = x$ when $x < 0$ and $G(x) = 1$ when $x > 2$). Then set $F = \chi_B$ where B is the set found in part c. Now $(F \circ G)^{-1}(\{1\}) = A$ is not Lebesgue measurable. So $F \circ G$ is not a Lebesgue measurable function.

The significance of this result is that we've proven that G is continuous but not Lebesgue measurable.

One more interesting observation Folland makes is that the collection of Borel sets $\mathcal{B}_{\mathbb{R}}$ only has the cardinality of the continuum, meaning that most measurable sets are not Borel.

To prove this, firstly note that by exercise 1.3 in my LaTeX math 240A notes (page 11), we know that $\mathrm{card}(\mathcal{B}_{\mathbb{R}}) \geq \mathfrak{c}$.

Also, consider the following lemmas:

1. **Proposition 0.14:**
   (a) If $\mathrm{card}(X) \leq \mathfrak{c}$ and $\mathrm{card}(Y) \leq \mathfrak{c}$, then $\mathrm{card}(X \times Y) \leq \mathfrak{c}$.

   Proof:
   It suffices to take $X = Y = \mathcal{P}(\mathbb{N})$ since then both X and Y have the largest cardinality we are allowing. Next, define $\psi, \phi : \mathbb{N} \to \mathbb{N}$ by $\psi(n) = 2n$ and $\phi(n) = 2n - 1$. Then $f : \mathcal{P}(\mathbb{N}) \times \mathcal{P}(\mathbb{N}) \to \mathcal{P}(\mathbb{N})$ given by $f(A, B) = \psi(A) \cup \phi(B)$ is a bijection.

   (b) If $\mathrm{card}(A) \leq \mathfrak{c}$ and $\mathrm{card}(\mathcal{E}_\alpha) \leq \mathfrak{c}$ for all $\alpha \in A$, then $\mathrm{card}(\bigcup_{\alpha \in A} \mathcal{E}_\alpha) \leq \mathfrak{c}$.

   Proof:
   For each $\alpha \in A$ there is a surjection $f_\alpha : \mathbb{R} \longrightarrow \mathcal{E}_\alpha$. So define the function $f : \mathbb{R} \times A \longrightarrow \bigcup_{\alpha \in A} \mathcal{E}_\alpha$ by $f(x, \alpha) = f_\alpha(x)$. This is a surjection. So, we know that $\mathrm{card}(\bigcup_{\alpha \in A} \mathcal{E}_\alpha) \leq \mathrm{card}(\mathbb{R} \times A)$ and the latter set by part (a) has no greater than the cardinality of the continuum.

2. If $\mathrm{card}(\mathcal{E}) \leq \mathfrak{c}$, then $\mathrm{card}(\mathcal{E}^\omega) \leq \mathfrak{c}$.

   To see this, we can assume $\mathcal{E} = \{0,1\}^\omega$ since then $\mathrm{card}(\mathcal{E}) = \mathfrak{c}$. Then note that we can use a diagonalization argument to create a bijection between between $\mathcal{E}$ and $\mathcal{E}^\omega$. Writing it out would be a pain so do it yourself.

Now recall from Folland's proposition 1.23 (the bonus proposition written on page 38 of my LaTeX notes for math 240A) the following construction of $\mathcal{B}_\mathbb{R}$.

Let S_Ω be a minimal uncountable set (by constructing S_Ω from $\mathbb{R}$ using the construction I copied from Munkres on page 14 of this pdf, we can guarentee that $S_\Omega \subseteq \mathbb{R}$ and thus $\mathrm{card}(S_\Omega) \leq \mathfrak{c}$).

Next, using 0 to refer to the least element of S_Ω, let $\mathcal{E}_0$ be the set of all rays of the form $[a, \infty)$ where $a \in \mathbb{R}$. Then for all other $\alpha \in S_\Omega$:

- If α has a direct predecessor β, then let $\mathcal{E}_\alpha$ be the collection of all countable unions of and complements of sets from $\mathcal{E}_\beta$.
- If α does not have a direct predecessor, then set $\mathcal{E}_\alpha = \bigcup_{\beta \in S_\alpha} \mathcal{E}_\alpha$.

Finally, $\mathcal{B}_\mathbb{R} = \bigcup_{\alpha \in S_\Omega} \mathcal{E}_\alpha$.

We obviously have that $\mathrm{card}(\mathcal{E}_0) = \mathfrak{c}$. Then using transfinite induction along with our two previously mentioned lemmas, we can conclude that $\mathrm{card}(\mathcal{E}_\alpha) \leq \mathfrak{c}$ for all $\alpha \in S_\Omega$. So by part (b) of proposition 0.14, we conclude that:

$$\mathrm{card}(\mathcal{B}_\mathbb{R}) = \mathrm{card}(\textstyle\bigcup_{\alpha \in S_\Omega} \mathcal{E}_\alpha) \leq \mathfrak{c}.$$

Since $\mathfrak{c} \leq \mathrm{card}(\mathcal{B}_\mathbb{R}) \leq \mathfrak{c}$, we know that $\mathrm{card}(\mathcal{B}_\mathbb{R}) = \mathfrak{c}$.

7/5/2025

I'm going to be taking more analysis notes from Folland. I'm starting with the section: <u>The Dual of $C_0(X)$</u>. Here, X refers to an LCH space.

To start out, we shall identify all positive bounded linear functionals on $C_0(X)$. Note that if I is such a functional on $C_0(X)$, then we know it is also a positive bounded linear functional on the subspace $C_c(X)$. Meanwhile going in reverse, we have that if $I(f) = \int f \mathrm{d}\mu$ is a positive linear function on $C_c(X)$ that is bounded, then we can uniquely extend it to a positive bounded linear functional on $C_0(X)$ by defining $I(f) = \lim_{n\to\infty} I(f_n)$ for any $f \in C_0(X)$ where $\{f_n\}_{n\in\mathbb{N}}$ is any sequence in $C_c(X)$ converging to f uniformly. So, given any Radon measure μ, we need to determine when $I(f) = \int f \mathrm{d}\mu$ is bounded.

Since X is open and μ is Radon, by the Riesz Representation theorem:

$$\mu(X) = \sup\{I(f) : f \in C_c(X),\ \mathrm{supp}(f) \subseteq X,\ 0 \leq f \leq 1\}.$$

The second condition is redundant and $I(f) = \int f \mathrm{d}\mu$. So we can rewrite this as $\mu(X) = \sup\{\int f \mathrm{d}\mu : f \in C_c(X),\ 0 \leq f \leq 1\}$. We now claim I is bounded iff $\mu(X) < \infty$, and that when I is bounded, $\|I\|_{\mathrm{op}} = \mu(X)$.

$(\Longrightarrow)$
Suppose $\mu(X) = \infty$. Then for any $N > 0$, there is a function $f \in C_c(X)$ with $0 \leq f \leq 1$ such that $\int f \mathrm{d}\mu \geq N$. And since $\|f\|_u \leq 1$, we know that if $|I(f)| \leq C\|f\|_u$, then $N \leq |\int f \mathrm{d}\mu| = |I(f)| \leq C$. This proves no finite C works for all $f \in C_c(X)$, and thus I is unbounded.

$(\Longleftarrow)$
Suppose $\mu(X) < \infty$ and then consider any $f \in C_c(X)$ with $\|f\|_u = 1$. Note that $|I(f)| = |\int f \mathrm{d}\mu| \leq \int |f| \mathrm{d}\mu$. Then since $0 \leq |f| \leq 1$, we know that $\int |f| \mathrm{d}\mu \leq \mu(X)$. So $\|I\|_{\mathrm{op}}$ exists and is at most $\mu(X)$.

To prove that $\mu(X) = \|I\|_{\mathrm{op}}$, let $\varepsilon > 0$ and pick $f \in C_c(X)$ with $0 \leq f \leq 1$ such that $\int f \mathrm{d}\mu > \mu(X) - \varepsilon$. Thus we have that $\|I\|_{\mathrm{op}}\|f\|_u > \mu(X) - \varepsilon$. Then since $\|f\|_u \leq 1$, we have that $\|I\|_{\mathrm{op}} > \mu(X) - \varepsilon$. Taking $\varepsilon \to 0$ finishes the proof.

So, the positive bounded linear functionals on $C_0(X)$ are precisely given by integration against finite Radon measures (and this correspondence is one-to-one by the Riesz Representation theorem). Next, we identify the other linear functionals on $C_0(X)$.

<u>Lemma 7.15:</u> If $I \in C_0(X, \mathbb{R})^*$, there exists positive functionals $I^{\pm} \in C_0(X, \mathbb{R})^*$ such that $I = I^+ - I^-$.

Proof:
If $f \in C_0(X, [0, \infty))$, define:

$$I^+(f) = \sup\{I(g) : g \in C_0(X, \mathbb{R}),\ 0 \leq g \leq f\}.$$

If $c \geq 0$, then clearly $I^+(cf) = cI^+(f)$. Meanwhile, let $f_1, f_2 \in C_0(X, [0, \infty))$. To show that $I^+(f_1 + f_2) = I^+(f_1) + I^+(f_2)$, first suppose $0 \leq g_1 \leq f_1$ and $0 \leq g_2 \leq f_2$. Then $0 \leq g_1 + g_2 \leq f_1 + f_2$. So, $I^+(f_1 + f_2) \geq I(g_1 + g_2) = I(g_1) + I(g_2)$. By taking $I(g_1) \to I^+(f_1)$ and $I(g_2) \to I^+(f_2)$, we then get that $I^+(f_1 + f_2) \geq I^+(f_1) + I^+(f_2)$.

On the other hand, if $0 \leq g \leq f_1 + f_2$, let $g_1 = \min(g, f_1)$ and $g_2 = g - g_1$. Thus $0 \leq g_1 \leq f_1$, $0 \leq g_2 \leq f_2$, and g_1, g_2 are continuous. This guarentees $g_1, g_2 \in C_0(X, [0, \infty))$. So, $I(g) = I(g_1) + I(g_2) \leq I^+(f_1) + I^+(f_2)$. And, taking $I(g) \to I^+(f_1 + f_2)$ gets us $I^+(f_1 + f_2) \leq I^+(f_1) + I^+(f_2)$.

Now, we extend I^+ to a positive linear functional in $C_0(X, \mathbb{R})^*$. Given $f \in C_0(X, \mathbb{R})$, let f^+ and f^- be the positive and negative parts of f. Then, $f^+, f^- \in C_0(X, [0, \infty))$. So, define $I^+(f) = I^+(f^+) - I^+(f^-)$. This is linear because if $c \in \mathbb{R}$, then ignoring the trivial edge case where $c = 0$:

$$I^+(cf) = \mathrm{sgn}(c)\,(I^+(|c|f^+) - I^+(|c|f^-))$$
$$= \mathrm{sgn}(c)|c|\,(I^+(f^+) - I^+(f^-)) = cI^+(f).$$

Also, suppose $f = g + h$ where $g, h \in C_0(X, \mathbb{R})$. Then $f^+ + g^- + h^- = f^- + g^+ + h^+$ where all the functions in that expression are in $C_0(X, [0, \infty))$. So, we know from our earlier work that:

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

Or in other words:

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

To show that I^+ is bounded, first note that if $f \in C_0(X, [0, \infty))$, then since $|I(g)| \leq \|I\|\|g\|_u \leq \|I\|\|f\|_u$ for all $0 \leq g \leq f$ and $I(0) = 0$, we have $0 \leq I^+(f) \leq \|I\|\|f\|_u$. (Note, this also proves I^+ is positive). Meanwhile, if $f \in C_0(X, \mathbb{R})$, then $I^+(f) = I^+(f^+) - I^+(f^-)$ where both terms in that difference are positive. Hence, we can say that:

$$|I^+(f)| \leq \max(I^+(f^+), I^-(f^-)) \leq \|I\| \max(\|f^+\|_u, \|f^-\|_u) = \|I\|\|f\|_u$$

Thus, we've finished constructing I^+. So now define $I^-(f) = I^+(f) - I(f)$. Then we know $I^- \in C_0(X, \mathbb{R})^*$ because $C_0(X, \mathbb{R})^*$ is a real vector space. Also, I^- is positive because if $f \geq 0$, then you can see from our definition of $I^+(f)$ on $C_0(X, [0, \infty))$ that $I^+(f) \geq I(f)$. Hence, $I^-(f) = I^+(f) - I(f)$ is also nonnegative. ■

Now any $I \in C_0(X)^*$ is uniquely determined by its restriction J to $C_0(X, \mathbb{R})$.

Why:
$I(f) = I(\mathrm{Re}(f) + i\mathrm{Im}(f)) = I(\mathrm{Re}(f)) + iI(\mathrm{Im}(f)) = J(\mathrm{Re}(f)) + iJ(\mathrm{Im}(f))$.

Next, there are two real linear functionals $J_1, J_2 \in C_0(X, \mathbb{R})^*$ such that $J = J_1 + iJ_2$. Specifically, set $J_1(f) = \mathrm{Re}(J(f))$ and $J_2 = \mathrm{Im}(J(f))$. Then clearly J_1 and J_2 are real linear functionals and they are bounded with $\|J_i\| \leq \|I\|$.

Using our lemma, we can decompose J_1 and J_2 into differences of positive bounded linear real functionals. I.e., we write $J = J_1^+ - J_1^- + i(J_2^+ - j_2^-)$.

Finally, define $I_1^+, I_1^-, I_2^+, I_2^-$ such that $I_1^+(f) = J_1^+(\mathrm{Re}(f)) + iJ_1^+(\mathrm{Im}(f))$ and the others have analogous definitions. Then all of our $I_i^\pm$ are well-defined complex linear functionals on $C_0(X)$ that are bounded since:

$$|I_i^\pm(f)| \leq \|J_i^\pm\|(\|\mathrm{Re}(f)\|_u + \|\mathrm{Im}(f)\|_u) \leq 2\|J_i^\pm\|\|f\|_u.$$

Also, all the $I_i^\pm$ are positive since if f is nonnegative, then $I_i^\pm(f) = J_i^\pm(f)$. This means that there are finite Radon measures $\mu_1, \mu_2, \mu_3, \mu_4$ such that $I_1^+(f) = \int f d\mu_1$, $I_1^-(f) = \int f d\mu_2$, $I_2^+(f) = \int f d\mu_3$. and $I_2^-(f) = \int f d\mu_4$.

Additionally:

$$\begin{aligned}
I(f) &= J(\mathrm{Re}(f)) + iJ(\mathrm{Im}(f)) \\
&= J_1(\mathrm{Re}(f)) + iJ_2(\mathrm{Re}(f)) + iJ_1(\mathrm{Im}(f)) + i^2 J_2(\mathrm{Im}(f)) \\
&= J_1^+(\mathrm{Re}(f)) - J_1^-(\mathrm{Re}(f)) + iJ_2^+(\mathrm{Re}(f)) - iJ_2^-(\mathrm{Re}(f)) \\
&\quad + iJ_1^+(\mathrm{Im}(f)) - iJ_1^-(\mathrm{Im}(f)) + i^2 J_2^+(\mathrm{Im}(f)) - i^2 J_2^-(\mathrm{Im}(f)) \\
&= I_1^+(f) - I_2^-(f) + iI_2^+(f) - iI_2^-(f)
\end{aligned}$$

So for any $I \in C_0(X)^*$, there are finite Radon measures $\mu_1, \mu_2, \mu_3, \mu_4$ such that $I(f) = \int f \mathrm{d}\mu$ where $\mu = \mu_1 - \mu_2 + i\mu_3 - i\mu_4$.

7/6/2025

I'm continuing on in Folland where I left off.

A signed Radon measure is a signed Borel measure such that it's positive and negative variations are Radon.

A complex Radon measure is a complex Borel measure such that it's real and imaginary variations are signed Radon measures.

Side note: Complex Borel measures are always finite on compact sets. Thus if X is an LCH space in which every open set is σ-compact, we know by theorem 7.8 that all complex Borel measures are Radon. In particular, if X is a second countable LCH space, then all complex Borel measures are Radon.

We denote the space of complex Radon measures on $(X, \mathcal{B}_X)$ as $M(X)$ and for $\mu \in M(X)$ we define $\|\mu\| = |\mu|(X)$ where $|\mu|$ is the total variation of μ.

Proposition 7.16: If μ is a complex Borel measure, then μ is Radon iff $|\mu|$ is Radon. Moreover, $M(X)$ is a vector space and $\mu \mapsto \|\mu\|$ is a norm on that space.

Proof:

By proposition 7.5 (which says that Radon measures are inner regular on all their σ-finite sets), we know that a finite positive measure $|\mu|$ is Radon iff for any Borel set E and $\varepsilon > 0$ there is an open set U and a compact set K with $K \subseteq E \subseteq U$ and $\mu(U - K) < \varepsilon$. From this we show the first assertion as follows. If $\mu = \mu_1 - \mu_2 + i\mu_3 - i\mu_4$ where all the μ_j are finite positive measures, and $|\mu|(U - K) < \varepsilon$, then $\mu_j(U - K) < \varepsilon$ for all j.

Why: (Also I'm going into more detail cause I am having trouble remembering how to work with the total variation of a complex measure.) Let ν be some positive measure with $\mu \ll \nu$. Then $\mu_j \ll \nu$ for each j, so for each j there are functions f_j with $\mathrm{d}\mu_j = f_j \mathrm{d}\nu$. Also, $\mathrm{d}\mu = (f_1 - f_2 + i(f_3 - f_4))\mathrm{d}\nu$ and $\mathrm{d}|\mu| = |f_1 - f_2 + i(f_3 - f_4)|\mathrm{d}\nu$.

Now since all the f_j are real-valued, we have:
$$|f_1 - f_2 + i(f_3 - f_4)| \geq \max(|f_1 - f_2|, |f_3 - f_4|).$$

Next, since $\mu_1 \perp \mu_2$ and $\mu_3 \perp \mu_4$ and all the measures are positive, we know that $\min(f_1, f_2) = 0$ and $\min(f_3, f_4) = 0$ ν-a.e. Hence,
$$\max(|f_1 - f_2|, |f_3 - f_4|) \geq \max(f_1, f_2, f_3, f_4) \text{ a.e.}$$

And so, we get $|\mu|(E) \geq \max(\mu_1(E), \mu_2(E), \mu_3(E), \mu_4(E))$ for all $E \in \mathcal{B}_X$.

Meanwhile if we can pick U_j, K_j for all j such that $\mu_j(U_j - K_j) < \varepsilon/4$, then set $U = \bigcap U_j$ and $K = \bigcup K_j$. Now, $|\mu|(U - K) < 4 \cdot \varepsilon/4 = \varepsilon$.

Why: By proposition 3.14,
$$\begin{aligned} |\mu| = |\mu_1 - \mu_2 + i\mu_3 - i\mu_4| &\leq |\mu_1| + |-\mu_2| + |i\mu_3| + |-i\mu_4| \\ &= \mu_1 + \mu_2 + \mu_3 + \mu_4. \end{aligned}$$

Then since $\mu_j(U) \leq \mu_j(U_j)$ and $\mu_j(K) \geq \mu(K_j)$ for all j, the claim holds.

Similar reasoning to that right above can show that $M(X)$ is closed under addition, and that $\|\mu_1 + \mu_2\| \leq \|\mu_1\| + \|\mu_2\|$. Also if $\mathrm{d}\mu = f\mathrm{d}\nu$ for some positive measure ν, then $c\mathrm{d}\mu = cf\mathrm{d}\nu$. So $|c\mathrm{d}\mu| = |c|\mathrm{d}|\mu|$, and from that it is clear that $|\mu|$ being Radon implies $|c\mathrm{d}\mu|$ is Radon. So, $M(X)$ is closed under scalar multiplication. Note this also shows that $\|c\mu\| = |c|\|\mu\|$ for all $c \in \mathbb{C}$ and $\mu \in M(X)$.

Finally, suppose $\mu \in M(X)$ with $\mu \neq 0$. Then there is some set $E \in \mathcal{B}_X$ such that $\mu(E) \neq 0$. Next $0 < |\mu(E)| \leq |\mu|(E)$ (see proposition 3.13). And since $|\mu|$ is Radon, we know that:
$$0 < |\mu|(E) = \inf\{|\mu|(U) : E \subseteq U \text{ where } U \text{ is open}\} \leq |\mu|(X) = \|\mu\|.$$

This proves, $M(X)$ is a normed vector space when equipped with $\|\cdot\|$. ■

7/7/2025

Before getting to the next theorem, I'd like to return to when I showed that for any $I \in C_0(X)^*$, there are finite Radon measures $\mu_1, \mu_2, \mu_3, \mu_4$ such that $I(f) = \int f\mathrm{d}\mu$ where $\mu = \mu_1 - \mu_2 + i\mu_3 - i\mu_4$.

A thing that Folland neglected to show is that while it's clear that $\mu_1 - \mu_2$ and $\mu_3 - \mu_4$ are the real and imaginary variations of μ respectively, it's not necessarily clear that μ_1 and μ_2 are the positive and negative variations respectively of $(\mu_1 - \mu_2)$ and likewise for μ_3 and μ_4 with respect to $(\mu_3 - \mu_4)$. So, I want to show that today since this will be relevant to the next propositions that Folland covers.

<u>Lemma (Riesz Representation theorem on $C_0(X,\mathbb{R})^*$):</u> There is a one-to-one correspondance between positive linear functionals in $C_0(X,\mathbb{R})^*$ and finite Radon measures on $(X,\mathcal{B}_X)$.

To start out, if $I \in C_0(X,\mathbb{R})^*$, then recall that there is a unique $J \in C_0(X)^*$ such that $J|_{C_0(x,\mathbb{R})} = I$. Namely, $J(f) = I(\mathrm{Re}(f)) + iI(\mathrm{Im}(f))$. Then I being positive means that J is positive. So by the Riesz Representation theorem, there is a unique finite Radon measure μ on $(X,\mathcal{B}_X)$ such that $J(f) = \int f\mathrm{d}\mu$. Then since $C_0(X,\mathbb{R}) \subseteq C_0(X)$, we have that $I(f) = \int f\mathrm{d}\mu$ for all $f \in C_0(X,\mathbb{R})$. At the same time, for any finite Radon measure μ, $f \mapsto \int f\mathrm{d}\mu$ is in $C_0(X,\mathbb{R})^*$. So, there is a bijective correspondence between finite Radon measures on X and $\{f \in C_0(X,\mathbb{R})^* : f \text{ is positive}\}$.

Now suppose $I \in C_0(X,\mathbb{R})^*$ and let $I = I^+ - I^-$ where $I^\pm \in C_0(X,\mathbb{R})^*$ are as we constructed in Lemma 7.15. As we just demonstrated, there are finite Radon measures μ_1 and μ_2 such that $I^+(f) = \int f\mathrm{d}\mu_1$ and $I^-(f) = \int f\mathrm{d}\mu_2$. In turn, setting $\mu = \mu_1 - \mu_2$ we have that $I(f) = \int f\mathrm{d}\mu$.

<u>Exercise 7.16:</u> The positive and negative variations of μ are the Radon measures μ_1 and μ_2 respectively.

Let μ^+ and μ^- be the positive and negative variations of μ, and let $E \in \mathcal{B}_X$ be a set such that $\mu^+(E) = 0$ and $\mu^-(E^\mathsf{C}) = 0$.

Fixing $f \in C_0(X,[0,\infty))$, note that:

$$\begin{aligned} I^+(f) &= \sup\{I(g) : g \in C_0(X,\mathbb{R}), 0 \le g \le f\} \\ &= \sup\{\textstyle\int g\mathrm{d}\mu^+ - \int g\mathrm{d}\mu^- : g \in C_0(X,\mathbb{R}), 0 \le g \le f\} \\ &\le \sup\{\textstyle\int g\mathrm{d}\mu^+ : g \in C_0(X,\mathbb{R}), 0 \le g \le f\} = \int f\mathrm{d}\mu^+ \end{aligned}$$

On the other hand, since μ_1,μ_2 are finite Radon measures and I showed yesterday that $M(X)$ is a vector space, I know that $\mu = \mu_1 - \mu_2$ is also a finite Radon measure. Also from yesterday, I know that that is equivalent to saying that $|\mu| = \mu^+ + \mu^-$ is Radon. Plus, μ being finite implies $|\mu|$ is finite. Hence, f vanishes outside of a set with finite measure (that set being all of X). So, for any $\varepsilon > 0$ we can apply Lusin's theorem to get a function $\phi \in C_c(X)$ with $\phi = f\chi_{E^\mathsf{C}}$ except on a set $F \in \mathcal{B}_X$ with $|\mu|(F) < \varepsilon$.

If we then set $\psi = \min(\mathrm{Re}(\phi)^+, f)$, we still have that $\psi = f\chi_{E^\mathsf{C}}$ except on F. But then also $\psi \in C_C(X,\mathbb{R}) \subseteq C_0(X,\mathbb{R})$ with $0 \le \psi \le f$. So:

$$\begin{aligned} I^+(f) \ge \textstyle\int \psi\mathrm{d}\mu &= \textstyle\int \psi\mathrm{d}\mu^+ - \int \psi\mathrm{d}\mu^- \\ &= \textstyle\int_F \psi\mathrm{d}\mu^+ + \int_{F^\mathsf{C}} f\chi_{E^\mathsf{C}}\mathrm{d}\mu^+ - \int_F \psi\mathrm{d}\mu^- - \int_{F^\mathsf{C}} f\chi_{E^\mathsf{C}}\mathrm{d}\mu^- \\ &\ge 0 + \textstyle\int_{F^\mathsf{C}} f\chi_{E^\mathsf{C}}\mathrm{d}\mu^+ - \int_F \psi\mathrm{d}\mu^- - \int_{F^\mathsf{C}} f\chi_{E^\mathsf{C}}\mathrm{d}\mu^- \\ &= \textstyle\int_{F^\mathsf{C}} f\mathrm{d}\mu^+ - \int_F \psi\mathrm{d}\mu^- - 0 \\ &= \textstyle\int f\mathrm{d}\mu^+ - \int_F f\mathrm{d}\mu^+ - \int_F \psi\mathrm{d}\mu^- - 0 \\ &\ge \textstyle\int f\mathrm{d}\mu^+ - 2\|f\|_u\mu(F) > \int f\mathrm{d}\mu^+ - 2\varepsilon\|f\|_u \end{aligned}$$

Since f was fixed, by taking $\varepsilon \to 0$ we have thus proven that $I^+(f) = \int f \mathrm{d}\mu^+$ for all $f \in C_0(X, [0,\infty))$. Then by considering positive and negative parts and making use of the linearity of both sides, we can easily see $I^+(f) = \int f \mathrm{d}\mu^+$ for all $f \in C_0(x, \mathbb{R})$. This proves that μ^+ is the unique Radon measure associated with I^+. Hence, $\mu^+ = \mu_1$.

Also, since $I^-(f) = I^+(f) - I(f) = \int f \mathrm{d}\mu^+ - (\int f \mathrm{d}\mu^+ - \int f \mathrm{d}\mu^-)$, we have $I^-(f) = \int f \mathrm{d}\mu^-$ for all $f \in C_0(X, \mathbb{R})$. So $\mu^- = \mu_2$. ■

Now as seen in the first lemma I showed today, if we extend $I^\pm \in C_0(X, \mathbb{R})^*$ to be a linear functional in $C_0(X)^*$, it doesn't change the measure μ at all. So, I'm done.

7/8/2025

Firstly, I'm going to finish describing $C_0(X)^*$.

Proposition 7.17 (The Riesz Representation Theorem): Let X be an LCH space, and for $\mu \in M(X)$ and $f \in C_0(X)$, let $I_\mu(f) = \int f \mathrm{d}\mu$. Then the map $\mu \mapsto I_\mu$ is an isometric isomorphism from $M(X)$ to $C_0(X)^*$.

Proof:
We already have shown that every $I \in C_0(X)^*$ is of the form I_μ for some $\mu \in M(X)$. On the other hand, if $\mu \in M(X)$, then we already know that I_μ is a linear function. Also, by proposition 3.13:

$$|\textstyle\int f \mathrm{d}\mu| \leq \int |f| \mathrm{d}|\mu| \leq \|f\|_u \|\mu\|.$$

So, I_μ is bounded with $\|I_\mu\| \leq \|\mu\|$.

All we have left to do is show $\|\mu\| \leq \|I_\mu\|$. So let $h = \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}$. Then since $|h| = 1$ by proposition 3.13 and $|\mu|$ is a finite Radon measure, we know by Lusin's theorem that for any $\varepsilon > 0$ there exists $f \in C_c(X)$ such that $\|f\|_u \leq \|h\|_u$ and $f = \overline{h}$ except on a set E with $|\mu|(E) < {}^{\varepsilon}/_{2}$. (Note that since $|\overline{h}| = 1$ almost everywhere, we have that $\|f\|_u = 1$.)

Now:

$$\begin{aligned}\|\mu\| = \textstyle\int 1 \mathrm{d}|\mu| &= \textstyle\int |h|^2 \mathrm{d}|\mu| \\ &= \textstyle\int \overline{h} \mathrm{d}\mu \leq |\int f \mathrm{d}\mu| + |\int (f - \overline{h}) \mathrm{d}\mu| \\ &\leq |I_\mu(f)| + \textstyle\int |f - \overline{h}| \mathrm{d}|\mu| \leq \|I_\mu\| + 2|\mu|(E) \\ &< \|I_\mu\| + \varepsilon\end{aligned}$$

Thus $\|\mu\| \leq \|I_\mu\|$ and we are done. ■

Corollary 7.18: If X is a compact Hausdorff space, then $C(X)^*$ is isometrically isomorphic to $M(X)$.

Next, I plan on taking a break from Folland chapter 7 in order to do some of the section 8.3 exercises in Folland that I never started or finished during the past Spring quarter.

Exercise 8.14 (Wirtinger's Inequality) If $f \in C^1([a,b])$ and $f(a) = f(b) = 0$, then:

$$\int_a^b |f(x)|^2 \mathrm{d}x \leq \left(\frac{b-a}{\pi}\right)^2 \int_a^b |f'(x)|^2 \mathrm{d}x$$

Hint: By a change of variables it suffices to assume $a = 0$ and $b = \frac{1}{2}$. Extend f To $[-\frac{1}{2}, \frac{1}{2}]$ by setting $f(-x) = -f(x)$, and then extend f to be periodic on $\mathbb{R}$. Check that f, thus extended, is in $C^1(\mathbb{T})$ and apply the Parseval identity.

Given our f, we can define $g(x) \coloneqq f(a + 2x(b-a))$. Then $g \in C^1([0, \frac{1}{2}])$ with $g(0) = g(\frac{1}{2}) = 0$ and $f(x) = g(\frac{x-a}{2(b-a)})$. Now suppose we prove the inequality for g. I.e., we show $\int_0^{1/2} |g(x)|^2 \mathrm{d}x \leq (\frac{1}{2\pi})^2 \int_0^{1/2} |g'(x)|^2 \mathrm{d}x$. Then:

- $\int_0^{1/2} |g(x)|^2 \mathrm{d}x = \int_0^{1/2} |f(a + 2x(b-a))|^2 \mathrm{d}x = 2(b-a) \int_a^b |f(y)|^2 \mathrm{d}y$,
- $(\frac{1}{2\pi})^2 \int_0^{1/2} |g'(x)|^2 \mathrm{d}x = (\frac{1}{2\pi})^2 \int_0^{1/2} |2(b-a) f'(a + 2x(b-a))|^2 \mathrm{d}x$
  $= (\frac{1}{2\pi})^2 (2(b-a))^3 \int_a^b |f'(y)|^2 \mathrm{d}y = 2(b-a)(\frac{b-a}{\pi})^2 \int_a^b |f'(y)|^2 \mathrm{d}y.$

By canceling out the $2(b-a)$ term (which is positive since $b > a$), we see the result still holds for f if it held for g.

But now we need to actually prove the result for g. To do this, extend out g to all of $\mathbb{R}$ by first setting $g(-x) \coloneqq -g(x)$ for $x \in [0, \frac{1}{2}]$, and then extending g to be periodic on $\mathbb{R}$. Note that this is well defined specifically because $g(0) = g(\frac{1}{2}) = 0$.

To see that g is in $C^1(\mathbb{T})$, note that since $g(x) = -g(-x)$ for $x \in (-\frac{1}{2}, 0)$ we have that $g'(x) = g'(-x)$ on $(-\frac{1}{2}, 0)$. Thus g is continuously differentiable on $(-\frac{1}{2}, 0)$ since we already know g is continuously differentiable on $(0, \frac{1}{2})$.

As for at $x = 0$, note that:

$$\lim_{h \to 0^-} \frac{g(h) - g(0)}{h} = \lim_{h \to 0^-} \frac{g(h)}{h} = \lim_{h \to 0^-} \frac{-g(-h)}{h} = \lim_{h \to 0^+} \frac{-g(h)}{-h} = \lim_{h \to 0^+} \frac{g(h)}{h} = \lim_{h \to 0^+} \frac{g(h) - g(0)}{h}$$

Thus $g'(0)$ still exists on the extended domain. Also, since $\lim_{x \to 0^+} g'(x) = g'(0)$, we know that $\lim_{x \to 0^-} g'(x) = \lim_{x \to 0^-} g'(-x) = \lim_{x \to 0^+} g'(x) = g'(0)$. So, g' is continuous at $t = 0$. Similar reasoning also works at $x = \frac{1}{2}$, although the looping structure of $\mathbb{T}$ makes the expressions slightly messier.

Now since $\mathbb{T}$ is compact, we know that $C(\mathbb{T}) \subseteq L^p$ for all p (and in particular, for $p = 2$). Thus both g and g' are in L^2. Applying Parseval's identity to g we get that:

$$\int_{-1/2}^{1/2} |g(x)|^2 dx = \|g\|^2_{L^2(\mathbb{T})} = \sum_{k \in \mathbb{Z}} |\widehat{g}(k)|^2 = \sum_{k \in \mathbb{Z}} |\int_{-1/2}^{1/2} g(x) e^{-2\pi i k x} \mathrm{d}x|^2$$

If we do integration by parts, then since $g(\frac{1}{2}) = g(-\frac{1}{2}) = 0$, we get for all $k \neq 0$ that:

$$\widehat{g}(k) = \int_{-1/2}^{1/2} g(x) e^{-2\pi i k x} \mathrm{d}x = \frac{1}{2\pi i k} \int_{-1/2}^{1/2} g'(x) e^{-2\pi i k x} \mathrm{d}x = \frac{1}{2\pi i k} \widehat{g'}(k)$$

Meanwhile, because of the way we extended g, we know g is an odd function. Thus, $\widehat{g}(0) = \int_{-1/2}^{1/2} g(x)\mathrm{d}x = 0$ and we've thus shown that:

$$\int_{-1/2}^{1/2} |g(x)|^2 dx \leq \sum_{\substack{k\in\mathbb{Z}\\k\neq 0}} \left|\frac{1}{2\pi i k}\widehat{g'}(k)\right|^2 = \frac{1}{4\pi^2}\sum_{\substack{k\in\mathbb{Z}\\k\neq 0}} \frac{1}{k^2}\left|\widehat{g'}(k)\right|^2$$

Now $\sum_{\substack{k\in\mathbb{Z}\\k\neq 0}} \frac{1}{k^2}\left|\widehat{g'}(k)\right|^2 \leq \sum_{k\in\mathbb{Z}} 1\cdot\left|\widehat{g'}(k)\right|^2$ and the latter is equal to $\|g'\|^2_{L^2(\mathbb{T})} = \int_{-1/2}^{1/2} |g'(x)|^2\mathrm{d}x$ by Parseval's identity.

Hence, we've proven that $\int_{-1/2}^{1/2} |g(x)|^2 dx \leq \left(\frac{1}{2\pi}\right)^2 \int_{-1/2}^{1/2} |g'(x)|^2\mathrm{d}x$.

Finally, since $|g(x)|^2$ and $|g'(x)|^2$ are both even on account of g being odd, we know that $\int_{-1/2}^{1/2} |g(x)|^2 dx = 2\int_0^{1/2} |g(x)|^2 dx$ and $\int_{-1/2}^{1/2} |g'(x)|^2\mathrm{d}x = 2\int_0^{1/2} |g'(x)|^2\mathrm{d}$. After canceling out the factor of 2, we've thus proven our desired inequality. ■

Exercise 8.16: Let $f_k = \chi_{[-1,1]} * \chi_{[-k,k]}$. (Also assume $k \in \mathbb{N}$ with $k > 0$).

(a) Compute $f_k(x)$ explicitely and show that $\|f\|_u = 2$.

You can fairly easily see that for any $x \in \mathbb{R}$, $f_k(x) = \int_{-k}^{k} \chi_{[-1,1]}(x-y)\mathrm{d}y$. Evaluating that gives the formula:

$$f_k(x) = \begin{cases} 2 & \text{if } |x| \leq k-1 \\ k-x+1 & \text{if } k-1 \leq x \leq k+1 \\ x+1+k & \text{if } -k-1 \leq x \leq -k+1 \\ 0 & \text{if } |x| \geq k+1 \end{cases}$$

From that it is hopefully clear that $\|f\|_u = 2$. After all, $f_k(0) = 2$. Also, $f_k(x) = \int_{-k}^{k} \chi_{[-1,1]}(x-y)\mathrm{d}y \leq \int \chi_{[-1,1]}(x-y)\mathrm{d}y = 2$.

(b) Show $f_k^\vee(x) = (\pi x)^{-2}\sin(2\pi x)\sin(2\pi k x)$, and $\|f_K^\vee\|_1 \to \infty$ as $k \to \infty$.

Recall from the homework that $\chi_{[-a,a]}^\wedge = \chi_{[-a,a]}^\vee = 2a\frac{\sin(2a\pi x)}{2\pi a x} = \frac{\sin(2\pi a x)}{\pi x}$.

Also, for any $f, g \in L^1$, by identical reasoning as we used to show $\widehat{f*g} = \widehat{f}\widehat{g}$, we know that $(f*g)^\vee = f^\vee g^\vee$. Therefore:

$$\begin{aligned} f_k^\vee(x) = \chi_{[-1,1]}^\vee(x)\chi_{[-k,k]}^\vee(x) &= \left(\frac{\sin(2\pi x)}{\pi x}\right)\left(\frac{\sin(2\pi k x)}{\pi x}\right) \\ &= (\pi x)^{-2}\sin(2\pi x)\sin(2\pi k x). \end{aligned}$$

Next, let $y = 2\pi k x$. Then:

$$\begin{aligned} \int |f_k^\vee(x)|\mathrm{d}x &= \int_{-\infty}^{\infty} |(\pi x)^{-2}\sin(2\pi x)\sin(2\pi k x)|\mathrm{d}x \\ &= \frac{1}{2\pi k}\int_{-\infty}^{\infty} \left|\frac{4k^2}{y^2}\sin\left(\frac{y}{k}\right)\sin(y)\right|\mathrm{d}y = \frac{2}{\pi}\int_{-\infty}^{\infty} \left|\frac{\sin(y/k)}{y/k}\cdot\frac{\sin(y)}{y}\right|\mathrm{d}y \\ &= \frac{4}{\pi}\int_0^{\infty} \left|\frac{\sin(y/k)}{y/k}\cdot\frac{\sin(y)}{y}\right|\mathrm{d}y. \end{aligned}$$

(the last equality holds because the integrand is even)

Now, because $\frac{\sin(x)}{x} \to 1$ as $x \to 0$, we know that $|\frac{\sin(y/k)}{y/k} \cdot \frac{\sin(y)}{y}|$ converges pointwise to $|\frac{\sin(y)}{y}|$ as $k \to \infty$. Also, observe that $|\frac{\sin(x)}{x}| \leq 1$ for all $x > 0$.

Proof:
Let $g(x) = \frac{\sin(x)}{x}$. Then clearly $|g(x)| \leq \frac{1}{x} \leq 1$ when $x \geq 1$.

Meanwhile, if $x < 1$, note that $g'(x) = \frac{x\cos(x)-\sin(x)}{x^2}$. Since $x^2 > 0$, it suffices to show that the numerator: $h(x) = x\cos(x) - \sin(x)$, is negative when $x < 1$ in order to prove that $g'(x)$ is not positive when $x < 1$, Luckily, note that $h(0) = 0$ and $h'(x) = -x\sin(x)$. Since $\sin(x) \geq 0$ for $x \leq \pi \approx 3.14$, we thus know that $h'(x) \leq 0$ for all $x \in [0,1]$. In turn, we know that $h(x) \leq h(0) = 0$ for all $x \in [0,1]$. So, we've proven that $g'(x)$ is not positive on $(0,1]$.

This proves that $g(x)$ is monotonically decreasing on $(0,1)$. And since $\lim_{x\to 0} g(x) = 1$, this proves that $g(x) \leq 1$ for all $x \in (0,1]$. Also, since $\sin(x) > 0$ when $0 < x < \pi \approx 3.14$, we know that $g(x) > 0$ for all $x \in (0,1]$. So, $|g(x)| \leq 1$ for all $x > 0$.

If we fix a constant $b > 0$, we have that: $|\frac{\sin(y/k)}{y/k} \cdot \frac{\sin(y)}{y} \cdot \chi_{[0,b]}(y)| \leq 1 \cdot 1 \cdot \chi_{[0,b]}(y)$ for all $k \in \mathbb{N}$. Hence by the dominated convergence theorem:

$$\liminf_{k\to\infty} \frac{4}{\pi}\int_0^\infty |\frac{\sin(y/k)}{y/k} \cdot \frac{\sin(y)}{y}|\mathrm{d}y \geq \lim_{k\to\infty} \frac{4}{\pi}\int_0^b |\frac{\sin(y/k)}{y/k} \cdot \frac{\sin(y)}{y}|\mathrm{d}y = \frac{4}{\pi}\int_0^b |\frac{\sin(y)}{y}|\mathrm{d}y$$

But now note that $\int_0^\infty |\frac{\sin(y)}{y}|\mathrm{d}y = \infty$.

$$\begin{aligned}\int_0^\infty |\frac{\sin(y)}{y}|\mathrm{d}y \geq \sum_{n=0}^{\infty}\int_{n\pi+\frac{\pi}{6}}^{n\pi+\frac{5\pi}{6}} |\frac{\sin(y)}{y}|\mathrm{d}y &= \sum_{n=0}^{\infty}\int_{n\pi+\frac{\pi}{6}}^{n\pi+\frac{5\pi}{6}} \frac{1}{2y}\mathrm{d}y \\ &\geq \frac{1}{2}\sum_{n=0}^{\infty}\int_{n\pi+\frac{\pi}{6}}^{n\pi+\frac{5\pi}{6}} \frac{1}{n\pi+\frac{5\pi}{6}}\mathrm{d}y \\ &= \frac{\pi}{3}\sum_{n=0}^{\infty}\frac{1}{n\pi+\frac{5\pi}{6}} \geq \frac{\pi}{3}\sum_{n=1}^{\infty}\frac{1}{n\pi} = \frac{1}{3}\sum_{n=1}^{\infty}\frac{1}{n} = \infty\end{aligned}$$

Thus, we can make $\int_0^b |\frac{\sin(y)}{y}|\mathrm{d}y$ arbitrarily big by making b big enough. Hence, we've proven that $\lim_{k\to\infty} \|f_k^\vee\|_1 = \lim_{k\to\infty} \frac{4}{\pi}\int_0^b |\frac{\sin(y)}{y}|\mathrm{d}y = \infty$.

Side note: while $\int_0^\infty \frac{\sin(y)}{y}\mathrm{d}y$ is not defined as a Lebesgue integral, it is defined as an improper Riemann integral and we can calculate that integral as follows.

Let $s > 0$. Then note that $\frac{\sin(y)}{y}$ and $e^{-sy}\chi_{[0,\infty)}$ are both in L^2. After all, $|\frac{\sin(y)}{y}|^2 \leq \chi_{[-1,1]}(y) + \frac{1}{y^2}\chi_{[-1,1]^\mathrm{c}}(y)$ and the right side is in L^2. Meanwhile, $\|e^{-sy}\|_2 = \frac{1}{2s}$. Thus by the Plancharel theorem, we know:

$$\int_0^\infty \frac{\sin(y)}{y}e^{-sy}\mathrm{d}y = \int_{-\infty}^\infty \mathcal{F}(\frac{\sin(y)}{y})\overline{\mathcal{F}(e^{-sy}\chi_{[0,\infty)}(y))}\mathrm{d}y$$

Now since $\chi_{[-a,a]}^{\vee} = \frac{\sin(2\pi a x)}{\pi x}$ for any $a \geq 0$, we can see that: $\mathcal{F}(\frac{\sin(y)}{y}) = \pi \chi_{[-\frac{1}{2\pi}, \frac{1}{2\pi}]}(\xi)$. Meanwhile:

$$\mathcal{F}(e^{-sy}\chi_{[0,\infty)}) = \int_0^\infty e^{-(s+2\pi i \xi)y}\mathrm{d}y = \frac{-1}{s+2\pi i \xi}(0-1) = \frac{1}{s+2\pi i \xi}$$

Hence, we've shown that:

$$\int_0^\infty \frac{\sin(y)}{y}e^{-sy}\mathrm{d}y = \pi \int_{-\frac{1}{2\pi}}^{\frac{1}{2\pi}} \overline{\left(\frac{1}{s+2\pi i \xi}\right)}\mathrm{d}\xi = \pi \int_{-\frac{1}{2\pi}}^{\frac{1}{2\pi}} \frac{s+2\pi i \xi}{s^2+4\pi^2\xi^2}\mathrm{d}\xi = \pi \int_{-\frac{1}{2\pi}}^{\frac{1}{2\pi}} \frac{s}{s^2+4\pi^2\xi^2}\mathrm{d}\xi$$

(Note, the last equality follows because we know that the imaginary part of the integral has to cancel since $\int_0^\infty \frac{\sin(y)}{y}e^{-sy}\mathrm{d}y$ is purely real-valued.)

Now:

$$\begin{aligned}\pi \int_{-\frac{1}{2\pi}}^{\frac{1}{2\pi}} \frac{s}{s^2+4\pi^2\xi^2}\mathrm{d}\xi &= \frac{s\pi}{4\pi^2}\int_{-\frac{1}{2\pi}}^{\frac{1}{2\pi}} \frac{1}{(\frac{s}{2\pi})^2+\xi^2}\mathrm{d}\xi \\ &= \frac{s}{4\pi}\left(\frac{2\pi}{s}\right)\left[\arctan(\tfrac{2\pi}{s}\xi)\right]_{\xi=-\frac{1}{2\pi}}^{\xi=\frac{1}{2\pi}} = \frac{1}{2}\left(\arctan(\tfrac{1}{s}) - \arctan(-\tfrac{1}{s})\right) = \arctan(\tfrac{1}{s})\end{aligned}$$

Thus, we've proven that $\int_0^\infty \frac{\sin(y)}{y}e^{-sy}\mathrm{d}y = \arctan(\frac{1}{s})$ for all $s > 0$.

Taking the limit as $s \to 0$, we get that $\int_0^\infty \frac{\sin(y)}{y}e^{-sy}\mathrm{d}y \to \frac{\pi}{2}$. That said, some care is needed since $\int_0^\infty \frac{\sin(y)}{y}\mathrm{d}y$ and $\lim_{s\to 0}\int_0^\infty \frac{\sin(y)}{y}e^{-sy}\mathrm{d}y$ are defined differently. In fact, we still have not showed that the former which is equal to $\lim_{b\to\infty}\int_0^b \frac{\sin(y)}{y}\mathrm{d}y$ exists. So, let's do that now.

Note that for any $b \in (0,\infty)$, there are unique $n \in \mathbb{Z}_{\geq 0}$ and $\alpha \in [0,\pi)$ such that $b = n\pi + \alpha$. Then for all $s \geq 0$, we have that:

$$\int_0^b \frac{\sin(y)}{y}e^{-sy}\mathrm{d}y = \sum_{j=0}^{n-1}(-1)^j\int_{j\pi}^{(j+1)\pi}\left|\frac{\sin(y)}{y}e^{-sy}\right|\mathrm{d}y + \int_{n\pi}^{n\pi+\alpha}\left|\frac{\sin(y)}{y}e^{-sy}\right|\mathrm{d}y$$

Now, the leftover term will approach 0 as $b \to \infty$ since it is at most $\frac{\alpha}{n\pi}$ when $b \geq 1$ and $n \to \infty$ as $b \to \infty$. Hence, letting $c_n = \int_{j\pi}^{(j+1)\pi}\left|\frac{\sin(y)}{y}e^{-sy}\right|\mathrm{d}y$, we know that: $\lim_{b\to\infty}\int_0^b \frac{\sin(y)}{y}e^{-sy}\mathrm{d}y = \sum_{n=0}^\infty (-1)^n c_n$. It's easily verified using the alternating series test that the series converges. This proves that our improper Riemann integral exists for all $s \geq 0$ (including $s = 0$).

Importantly, this series also converges uniformly over all $s \in [0,\infty)$. To see why, observe that since $(c_n)_{n\in\mathbb{N}}$ is a strictly decreasing sequence, for any $N \geq 0$ we have that: $c_N \geq |\sum_{n=N}^\infty (-1)^n a_n|$. This can be proven via induction fairly easily. Next, making s larger makes all the c_n strictly smaller. So, by picking N large enough so that $c_N < \varepsilon$ when $s = 0$, we can guarentee that $c_N < \varepsilon$ for all s. It then follows that the error from the limit point: $|\sum_{n=N}^\infty (-1)^n c_n|$, is also less than ε for all s.

With that, we know there is some $b > 0$ such that:

$$\left|\int_0^\infty \frac{\sin(y)}{y}e^{-sy}\mathrm{d}y - \int_0^b \frac{\sin(y)}{y}e^{-sy}\mathrm{d}y\right| < {}^{\varepsilon}/_{4} \text{ for all } s \geq 0.$$

Also, by dominated convergence theorem (its 4am and I don't want to type out verifications for all the conditions), we know that $\int_0^b \frac{\sin(y)}{y} e^{-sy} \mathrm{d}y \to \int_0^b \frac{\sin(y)}{y} \mathrm{d}y$ as $s \to 0$. So, there is some $s > 0$ such that:

$$\left| \int_0^b \frac{\sin(y)}{y} \mathrm{d}y - \int_0^b \frac{\sin(y)}{y} e^{-sy} \mathrm{d}y \right| < \varepsilon/4.$$

Also, by making s potentially smaller, we can also guarentee that:

$$\left| \frac{\pi}{2} - \int_0^\infty \frac{\sin(y)}{y} e^{-sy} \mathrm{d}y \right| < \varepsilon/4.$$

And chaining those together, we get that:

$$\begin{aligned} \left| \int_0^\infty \frac{\sin(y)}{y} \mathrm{d}y - \frac{\pi}{2} \right| \leq & \left| \int_0^\infty \frac{\sin(y)}{y} \mathrm{d}y - \int_0^b \frac{\sin(y)}{y} \mathrm{d}y \right| + \left| \int_0^b \frac{\sin(y)}{y} \mathrm{d}y - \int_0^b \frac{\sin(y)}{y} e^{-sy} \mathrm{d}y \right| \\ & + \left| \int_0^b \frac{\sin(y)}{y} e^{-sy} \mathrm{d}y - \int_0^\infty \frac{\sin(y)}{y} e^{-sy} \mathrm{d}y \right| + \left| \int_0^\infty \frac{\sin(y)}{y} e^{-sy} - \frac{\pi}{2} \right| \\ & < \varepsilon/4 + \varepsilon/4 + \varepsilon/4 + \varepsilon/4 = \varepsilon \end{aligned}$$

Taking $\varepsilon \to 0$, we've finally shown that $\lim\limits_{b \to \infty} \int_0^b \frac{\sin(y)}{y} \mathrm{d}y = \frac{\pi}{2}$.

(c) Prove that $\mathcal{F}(L^1)$ is a proper subset of C_0.

To start off with, recall that if $f \in L^1$ and $\widehat{f} = 0$, then $f = 0$ a.e. As a corollary to this, we have that if $f, g \in L^1$ and $\widehat{f} = \widehat{g}$, then $f = g$ a.e. This is because $(f - g)^\wedge = 0$ implies that $f - g = 0$ a.e. So, we know that $\mathcal{F}$ is an injective map from L^1 to C_0. If $\mathcal{F}$ was also surjective, then we would know that $\mathcal{F}$ is a bijection, and that therefore a function $\mathcal{F}^{-1} : C_0 \to L^1$ exists. Also, by the open map theorem, we would know that $\mathcal{F}^{-1}$ is bounded.

However, in part (b) we found that $\|f_k\|_u = 2$ for all $k \in \mathbb{N}$ but $\|f_k^\vee\|_1 \to \infty$ as $k \to \infty$. Importantly, we can see from our work earlier that $f_k^\wedge = f_k^\vee$ and $\|f_k^\vee\|_1 < \infty$ for all k. After all, $f_k^\vee(y)$ is bounded by 1 when $|y| \leq 1$ and by $\frac{k}{y^2}$ when $|y| \geq 1$. So by the Fourier inversion theorem, we know that $(f_k^\vee)^\wedge = f_k$ (with equality holding everywhere since both sides are continuous). And so, $\mathcal{F}^{-1}(f_k) = f_k^\vee$.

This proves that $\mathcal{F}^{-1}$ is not bounded since $\|\mathcal{F}^{-1}(f_k)\|_1$ can be made arbitrarily large even while $\|f_k\|_u = 2$ for all k.

7/10/2025

Today I'm gonna do more problems from chapter 8 of Folland.

Recall that for $f \in L^p(\mathbb{R})$, if there exists $h \in L^p(\mathbb{R})$ such that $\lim_{y \to 0} \|y^{-1}(\tau_{-y} f - f) - h\|_p = 0$, we call h the (strong) L^p derivative of f. If $f \in L^p(\mathbb{R}^n)$, L^p partial derivatives of f are defined similarly. (Also, the notation $\tau_y f(x)$ refers to $f(x - y)$.)

Exercise 8.8: Suppose that p and q are conjugate exponents, $f \in L^p$, $g \in L^q$, and the L^p derivative $\partial_j f$ exists. Then $\partial_j(f * g)$ exists (in the ordinary sense) and equals $(\partial_j f) * g$.

Note that:

$$\lim_{t \to 0} \frac{(f*g)(x+te_j) - (f*g)(x)}{t} = ((\partial_j f) * g)(x) \text{ iff } \lim_{t \to 0} \left| \frac{(f*g)(x+te_j) - (f*g)(x)}{t} - ((\partial_j f) * g)(x) \right| = 0.$$

Now for all $t \neq 0$:

$$0 \leq \left| \frac{(f*g)(x+te_j)-(f*g)(x)}{t} - ((\partial_j f) * g)(x) \right|$$
$$= \left|t^{-1} \int \left(f(x+te_j-y) - f(x-y)\right) g(y)\mathrm{d}y - \int \partial_j f(x-y)g(y)\mathrm{d}y\right|$$
$$= \left| \int t^{-1}(\tau_{-te_j} f - f)(x-y)g(y)\mathrm{d}y - \int \partial_j f(x-y)g(y)\mathrm{d}y \right|$$
$$\leq \int |t^{-1}(\tau_{-te_j} f - f)(x-y) - \partial_j f(x-y)||g(y)|\mathrm{d}y$$
$$\leq \|t^{-1}(\tau_{-te_j} f - f) - \partial_j f\|_p \|g\|_q$$

Since $\|g\|_q$ is fixed and $\|t^{-1}(\tau_{-te_j} f - f) - \partial_j f\|_p \to 0$ as $t \to 0$, we've thus shown that $\left| \frac{(f*g)(x+te_j)-(f*g)(x)}{t} - ((\partial_j f) * g)(x) \right| \to 0$ as $t \to 0$.

Exercise 8.9: Let $1 \leq p < \infty$. If $f \in L^p(\mathbb{R})$, the L^p derivative of f (call it h; see Exercise 8) exists iff f is absolutely continuous on every bounded interval (perhaps after modification on a null set) and its pointwise derivative f' is in L^p, in which case $h = f'$ a.e.

($\Longrightarrow$)

Suppose f has an L^p derivative h. Then setting $\varphi(x) = (1 - |x|)\chi_{[-1,1]}$, note that $\int \varphi(x)\mathrm{d}x = 1$ and $0 \leq \varphi \leq 1 \leq \frac{4}{(1+|x|)^2}$. Thus, φ satisfies the hypothesis of theorem 8.15 (see page 31 of my paper notes) and so we know that:

- $(f * \varphi_{1/n})(x) = \int f(x-y) \cdot n\varphi(ny)\mathrm{d}y \to f(x)$ as $n \to \infty$ for all $x \in L_f$,
- $(h * \varphi_{1/n})(x) = \int h(x-y) \cdot n\varphi(ny)\mathrm{d}y \to h(x)$ as $n \to \infty$ for all $x \in L_h$

(where L_f and L_h are the Lebesgue sets of f and h respectively).

Side note: If $g \in L^1_{\text{loc}}$, then we know that $(L_g)^{\mathrm{C}}$ has measure zero. Now while it's obvious that $L^1, L^\infty \subseteq L^1_{\text{loc}}$, I'm currently realizing I've never justified to myself why $L^p \subseteq L^1_{\text{loc}}$ for all $1 < p < \infty$.

If E is a measurable set with finite measure, then the measure restricted to E does not have any sets of arbitrarily large measure. Thus for any $p, q \in (0, \infty)$ with $p < q$, we have that $L^q(E) \subseteq L^p(E)$. Specifically, this means that for any $1 < p < \infty$, $L^p(E) \subseteq L^1(E)$. It follows that $L^p \subseteq L^1_{\text{loc}}$ for all $1 < p < \infty$.

Also, if q is the conjugate exponent of p, we know that $\phi_{1/n} \in L^q$. Therefore, by the previous exercise we know that $(f * \phi_{1/n})' = h * \varphi_{1/n}$. Additionally:

$$|h * \varphi_{1/n}(x)| \leq \int |h(x-y)\varphi_{1/n}(y)|\mathrm{d}y \leq \|h\|_p \left(\int_{-1/n}^{1/n} |n(1-|nx|)|^q \mathrm{d}x\right)^{1/q}$$
$$\leq n\|h\|_p \left(\int_{-1/n}^{1/n} 1\mathrm{d}x\right)^{1/q} = n\left(\tfrac{2}{n}\right)^{1/q} \|h\|_p \leq 2^{1/q}\|h\|_p$$

This tells us that for all $n \in \mathbb{N}$, $f * \varphi_{1/n}$ has a bounded derivative. It then follows by the mean value theorem that $f * \varphi_{1/n}$ is absolutely continuous. So for any $a, x \in \mathbb{R}$ with $a < x$, we have that:

$$(f * \varphi_{1/n})(x) - (f * \varphi_{1/n})(a) = \int_a^x (f * \varphi_{1/n})'(y)\mathrm{d}y = \int_a^x (h * \varphi_{1/n})(y)\mathrm{d}y$$

And, since $h * \varphi_{1/n} \to h$ pointwise a.e. and $2^{1/q}\|h\|_p\chi_{[a,x]} \in L^1$, we know by dominated convergence theorem that:

$$\lim_{n\to\infty}((f*\varphi_{1/n})(x) - (f*\varphi_{1/n})(a)) = \lim_{n\to\infty}\int_a^x (h*\varphi_{1/n})(y)\mathrm{d}y = \int_a^x h(y)\mathrm{d}y$$

Now, we're finally ready to show the right hand side of our implication. Suppose $a, b \in L_f$ are fixed with $a < b$. Then for any $x \in L_f \cap [a,b]$, we have that:

$$f(x) - f(a) = \lim_{n\to\infty}((f*\varphi_{1/n})(x)) - \lim_{n\to\infty}((f*\varphi_{1/n})(a)) = \int_a^x h(y)\mathrm{d}y$$

By redefining f on the null space $(L_f)^{\mathsf{C}} \cap [a,b]$, we can thus guarentee that $f(x) - f(a) = \int_a^x h(y)\mathrm{d}y$ for all $x \in [a,b]$. In turn, by the fundamental theorem of calculus we know f is absolutely continuous on $[a,b]$ and that $h = f'$ a.e. on $[a,b]$.

If $I \subseteq \mathbb{R}$ is any arbitrary bounded interval, then we can still apply the former reasoning by finding $a, b \in L_f$ such that $I \subseteq [a,b]$. Then f being absolutely continuous on $[a,b]$ implies that f is absolutely continuous I. Also, since $\mathbb{R}$ can be completed covered by these intervals, we know that $f' = h$ a.e. The only snag we still have to sort out is to show that our redefinitions of $f(x)$ for $x \in (L_f)^{\mathsf{C}}$ are well defined (i.e. not dependent on our choice of $a, b \in L^f$.)

> Suppose $a_1, a_2 \in L_f$ and without loss of generality assume $a_1 < a_2 < x$. Then:
> $$\left(\int_{a_1}^x h(y)\mathrm{d}y + f(a_1)\right) - \left(\int_{a_2}^x h(y)\mathrm{d}y + f(a_2)\right) = \int_{a_1}^{a_2} h(y)\mathrm{d}y - (f(a_2) - f(a_1)).$$
>
> Since $a_1, a_2 \in L^f$, we know that $\int_{a_1}^{a_2} h(y)\mathrm{d}y = f(a_2) - f(a_1)$. So our above expression equals 0 and we've shown that:
> $$f(x) = \int_{a_1}^x h(y)\mathrm{d}y + f(a_1) = \int_{a_2}^x h(y)\mathrm{d}y + f(a_2) \text{ is well defined.}$$

($\Longleftarrow$)

Note that if $y > 0$, then our assumptions about f tell us that:

$$\begin{aligned}\frac{f(x+y)-f(x)}{y} - f'(x) = \frac{1}{y}\int_x^{x+y} f'(t)\mathrm{d}t - f'(x) &= \frac{1}{y}\int_x^{x+y} f'(t) - f'(x)\mathrm{d}t \\ &= \frac{1}{y}\int_0^y f'(x+t) - f'(x)\mathrm{d}t\end{aligned}$$

Similarly, if $y < 0$, then we know:

$$\begin{aligned}\frac{f(x+y)-f(x)}{y} - f'(x) = \frac{-1}{y}\int_{x+y}^x f'(t)\mathrm{d}t - f'(x) &= \frac{-1}{y}\int_{x+y}^x f'(t) - f'(x)\mathrm{d}t \\ &= \frac{-1}{y}\int_y^0 f'(x+t) - f'(x)\mathrm{d}t\end{aligned}$$

In either case, we can see that:

$$\left|\frac{f(x+y)-f(x)}{y} - f'(x)\right| \leq \int_{-|y|}^{|y|}\frac{1}{|y|}|\tau_{-t}f'(x) - f'(x)|\mathrm{d}t$$

Thus by Minkowski's inequality for integrals:

$$\left\|\frac{f(x+y)-f(x)}{y} - f'(x)\right\|_p \leq \left\|\int_{-|y|}^{|y|}\frac{1}{|y|}|\tau_{-t}f'(x) - f'(x)|\mathrm{d}t\right\|_p \leq \frac{1}{|y|}\int_{-|y|}^{|y|}\|\tau_{-t}f'(x) - f'(x)\|_p\mathrm{d}t$$

And since translation is continuous with respect to the L^p norm for $1 \leq p < \infty$, we know that $\|\tau_{-t}f'(x) - f'(x)\|_p \to 0$ as $t \to 0$. Hence given $\varepsilon > 0$, we have for $|y|$ small enough that:

$$\frac{1}{|y|}\int_{-|y|}^{|y|}\|\tau_{-t}f'(x) - f'(x)\|_p\mathrm{d}t < \frac{1}{|y|}\int_{-|y|}^{|y|}\varepsilon\mathrm{d}t = \frac{2|y|\varepsilon}{|y|} = 2\varepsilon$$

By taking $\varepsilon \to 0$, this proves that $\left\|\frac{f(x+y)-f(x)}{y} - f'(x)\right\|_p \to 0$ as $y \to 0$. Hence f' is an L^p derivative of f. ■

So what's the significance of this result?

- A function on $\mathbb{R}$ having an L^p derivative is a strictly stronger assumption than the function just being differentiable almost everywhere.
- Any two L^p derivatives of a function are equal a.e. to the ordinary derivative of the function. Thus there's at most one L^p derivative of any function in $L^p(\mathbb{R})$.
- Any function $L^p(\mathbb{R})$ that is differentiable a.e. and whose derivative is bounded and also in L^p has an L^p derivative.

7/11/2025

Exercise 8.18: Suppose $f \in L^2(\mathbb{R})$.

(a) The L^2 derivative f' exists iff $\xi\widehat{f} \in L^2$, in which case $\widehat{f'}(\xi) = 2\pi i\xi\widehat{f}(\xi)$.

($\Longrightarrow$)
Once again set $\varphi(x) = (1-|x|)\chi_{[-1,1]}$. Then by theorem 8.14(a) (see page 29 of my paper notes): $f * \varphi_{1/n} \to f$ in L^2 as $n \to \infty$. In turn, since the Fourier transform is continuous on L^2, we know that $(f * \varphi_{1/n})^\wedge \to \widehat{f}$ as $n \to \infty$.

Next, note that $f * \varphi_{1/n} \in C^1$.

Why: Recall from exercise 8.8 that $(f * \varphi_{1/n})' = f' * \varphi_{1/n}$. Also, since $f' \in L^1_{\text{loc}}$ and $\varphi_{1/n} \in C^0$ has compact support, we know from exercise 8.7 (which was a homework problem in Math 240C), that $f' * \varphi_{1/n} \in C^0$.

Also, since f, f' and $\varphi_{1/n}$ are all in L^2, we know by proposition 8.8 (see page 25 of my paper notes) that $f * \varphi_{1/n} \in C_0$, and we know by Young's inequality (see page 26 of my paper notes) that $f * \varphi_{1/n}, f' * \varphi_{1/n} \in L^1$. All together, this lets us conclude via integration by parts that:

$$(f * \varphi_{1/n})^\wedge = \tfrac{1}{2\pi i\xi}((f * \varphi_{1/n})')^\wedge = \tfrac{1}{2\pi i\xi}(f' * \varphi_{1/n})^\wedge.$$

Finally, since $f' * \varphi_{1/n} \to f'$ in L^2 as $n \to \infty$ and the Fourier transform is continuous on L^2, we know that:

$$\widehat{f}(\xi) = \lim_{n\to\infty}(f * \varphi_{1/n})^\wedge(\xi) = \tfrac{1}{2\pi i\xi}\lim_{n\to\infty}(f' * \varphi_{1/n})^\wedge(\xi) = \tfrac{1}{2\pi i\xi}\widehat{f'}(\xi) \text{ a.e.}$$

Since $\frac{1}{2\pi i}\widehat{f'}$ is in L^2, this thus proves that $\xi\widehat{f} \in L^2$. Also, by rearranging out expression we get that $\widehat{f'} = 2\pi i\xi\widehat{f}(\xi)$.

($\Longleftarrow$)
Define $h(\xi) = 2\pi i\xi\widehat{f}(\xi)$. Then by assumption we know that $h \in L^2$. So, there exists a function $H \in L^2$ such that $\widehat{H} = h$. And since the Fourier transform is a continuous isometric linear operator on L^2, we know that for all $y \neq 0$:

$$\begin{aligned}\|\tfrac{1}{y}(\tau_{-y}f - f) - H\|_2 &= \|\mathcal{F}(\tfrac{1}{y}(\tau_{-y}f - f) - H)\|_2 \\ &= \|\tfrac{1}{y}(\mathcal{F}(\tau_{-y}f) - \mathcal{F}(f)) - h\|_2\end{aligned}$$

Now we claim that $\mathcal{F}(\tau_{-y}f) = e^{2\pi i\xi y}\mathcal{F}(f)$ for all $f \in L^2$.

Proof:
Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of Schwartz functions converging to $f \in L^2$. Then since $\int g = \int \tau_{-y}g$ for all functions g, we can easily see that $\tau_{-y}f_n \to \tau_{-y}f$ in L^2. Therefore, $\mathcal{F}(\tau_{-y}f) = \lim_{n\to\infty} \mathcal{F}(\tau_{-y}f_n)$.

Next, since $f_n \in L^1$ for all n, we know that $\mathcal{F}(\tau_{-y}f_n)(\xi) = e^{2\pi i\xi y}\widehat{f_n}(\xi)$. Then finally, since the Fourier transform is continuous on L^2, we have that $\widehat{f_n} \to \widehat{f}$ in L^2 as $n \to \infty$. By passing to a subsequence, we can assume $\widehat{f_n} \to \widehat{f}$ pointwise a.e. And so, $\mathcal{F}(\tau_{-y}f) = \lim_{n\to\infty} e^{2\pi i\xi y}\widehat{f_n}(\xi) = e^{2\pi i\xi y}\widehat{f}(\xi)$ a.e.

Thus, we know that:

$$\begin{aligned}\left|\tfrac{1}{y}(\mathcal{F}(\tau_{-y}f) - \mathcal{F}(f)) - h\right|^2 &= \left|\left(\tfrac{1}{y}e^{2\pi i\xi y} - \tfrac{1}{y} - 2\pi i\xi\right)\widehat{f}(\xi)\right|^2 \\ &= \left|\left((\tfrac{\cos(2\pi\xi y)}{y} - \tfrac{1}{y}) + i(\tfrac{\sin(2\pi\xi y)}{y} - 2\pi\xi)\right)\widehat{f}(\xi)\right|^2 \\ &= \left|\left((\tfrac{\cos(2\pi\xi y)-1}{\xi y}) + i(\tfrac{\sin(2\pi\xi y)}{y\xi} - 2\pi)\right)\xi\widehat{f}(\xi)\right|^2\end{aligned}$$

Now, note that $\lim_{y\to 0}\frac{\cos(2\pi\xi y)-1}{\xi y} = 2\pi\lim_{t\to 0}\frac{\cos(t)-1}{t} = 2\pi\cdot 0 = 0$ for all $\xi \neq 0$. Similarly, we have that $\lim_{y\to 0}\frac{\sin(2\pi\xi y)}{y\xi} = 2\pi\lim_{t\to 0}\frac{\sin(t)}{t} = 2\pi\cdot 1$ for all $\xi \neq 0$. This proves that $|\frac{1}{y}(\mathcal{F}(\tau_{-y}f) - \mathcal{F}(f)) - h|^2 \to 0$ pointwise a.e. as $y \to 0$.

Meanwhile, note that $\left|\frac{\sin(2\pi x)}{x}\right|$ and $\left|\frac{\cos(2\pi x)-1}{x}\right|$ are both less than or equal to 2π on their domains. Therefore, we can get that for all $\xi \neq 0$ and $y \neq 0$, we have that:

$$|(\tfrac{\cos(2\pi\xi y)-1}{\xi y}) + i(\tfrac{\sin(2\pi\xi y)}{y\xi} - 2\pi)| \leq |(\tfrac{\cos(2\pi\xi y)-1}{\xi y})| + |(\tfrac{\sin(2\pi\xi y)}{y\xi} - 2\pi)| \leq 6\pi$$

(I'm not sure how to prove $|\frac{\cos(2\pi x)-1}{x}| \leq 2\pi$ without pulling out numerical methods. But you'll see that it is true if you graph it.)

Thus using $36\pi^2|\xi\widehat{f}(\xi)|^2$ as our upper bound function (which is in L^1 since $\xi\widehat{f} \in L^2$), we can conclude via the dominated convergence theorem that:

$$\lim_{y\to 0}\|\tfrac{1}{y}(\mathcal{F}(\tau_{-y}f) - \mathcal{F}(f)) - h\|_2^2 = \lim_{y\to 0}\int\left|\tfrac{1}{y}(\mathcal{F}(\tau_{-y}f) - \mathcal{F}(f)) - h\right|^2 = 0$$

So, f has $H = h^\vee$ as it's L^2 derivative.

7/12/2025

Ok. I think that in order to prove part (b) of exercise 8.18, I need to make a pit stop in the exercises of section 3.5 of Folland. This is because Folland's hinted solution route is to use integration by parts. However, right now I've only shown that you can do integration by parts if the two functions in your integrand are continuously differentiable. Yet that's not guarenteeable in exercise 8.18(b). So, my current objective is to weaken my requirements for doing integration by parts.

Exercise 3.35: If F and G are absolutely continuous on $[a,b]$, then so is FG and:
$$\int_a^b (FG' + GF')(x)\mathrm{d}x = F(b)G(b) - F(a)G(a)$$

Proof:
By extreme value theorem, there exists $M \geq 0$ such that $\max(|F(x)|, |G(x)|) \leq M$ for all $x \in [a,b]$. Now for any $\varepsilon > 0$, let $\delta > 0$ be such that for all finite collections of disjoint intervals $(a_1, b_1), \ldots, (a_n, b_n) \subseteq [a,b]$ with $\sum_{i=1}^n |b_i - a_i| < \delta$, we have:
$$\sum_{i=1}^n |F(b_i) - F(a_i)| < \tfrac{\varepsilon}{2M} \text{ and } \sum_{i=1}^n |G(b_i) - G(a_i)| < \tfrac{\varepsilon}{2M}$$

Then we have:
$$\begin{aligned}
\sum_{i=1}^n |F(b_i)G(b_i) - F(a_i)G(a_i)| &\leq \sum_{i=1}^n |F(b_i)G(b_i) - F(b_i)G(a_i)| + \sum_{i=1}^n |F(b_i)G(a_i) - F(a_i)G(a_i)| \\
&= \sum_{i=1}^n |F(b_i)||G(b_i) - G(a_i)| + \sum_{i=1}^n |G(a_i)||F(b_i) - F(a_i)| \\
&\leq M \sum_{i=1}^n |G(b_i) - G(a_i)| + M \sum_{i=1}^n |F(b_i) - F(a_i)| \\
&< M\tfrac{\varepsilon}{2M} + M\tfrac{\varepsilon}{2M} = \varepsilon
\end{aligned}$$

Hence, FG is also absolutely continuous on $[a,b]$. It follows by the fundamental theorem of calculus for Lebesgue integrals that:
$$F(b)G(b) - G(b)G(a) = \int_a^b (FG)'(x)\mathrm{d}x = \int_a^b (FG' + GF')(x)\mathrm{d}x. \blacksquare$$

The following is also tangentially relevant to exercise 8.18(b) in addition to being interesting in its own right. A function $f : \mathbb{R} \to \mathbb{C}$ is <u>singular</u> if f is continuous everywhere, f' exists a.e. with $f' = 0$ a.e., and f is not a constant function.

Recalling page 53 of this journal, it's easy to see that the Cantor function is singular (if you continuously extend it to be constant outside $[0,1]$). However, it's also constant on a small enough neighborhood around almost every point. Hence, it doesn't defy our intution too overly much. In the next two exercises however, we'll construct a strictly increasing singular function.

Exercise 3.39: If $(F_k)_{k\in\mathbb{N}}$ is a sequence of nonnegative increasing functions on $[a,b]$ such that $F(x) = \sum_{k=1}^\infty F_k(x) < \infty$ for all $x \in [a,b]$, then $F'(x) = \sum_{k=1}^\infty F_k'(x)$ for a.e. $x \in [a,b]$.

For all k, define:
$$G_k(x) = \begin{cases} F_k(a+) & \text{if } x \leq a \\ F_k(x+) & \text{if } x \in [a,b) \\ F_k(b) & \text{if } x \geq b \end{cases} \quad \text{and} \quad G(x) = \begin{cases} F(a+) & \text{if } x \leq a \\ F(x+) & \text{if } x \in [a,b) \\ F(b) & \text{if } x \geq b \end{cases}$$

Then by theorem 3.23, we know that $G_k' = F_k'$; $G' = F'$ a.e. on $[a,b]$ Also, each G_k is a nonnegative monotone increasing function with $G(x) = \sum_{k=1}^{\infty} G_k(x) < \infty$ for all $x \in \mathbb{R}$.

Why: For any $x \in [a,b)$, we can apply the dominated convergence theorem to $l^1(\mathbb{N})$ using the upper bound $F(b) = \sum_{k=1}^{\infty} F_k(b)$ in order to get that:

$$G(x) = F(x+) = \lim_{t \to 0^+} \sum_{k=1}^{\infty} F_k(x+t) = \sum_{k=1}^{\infty} \lim_{t \to 0^+} F_k(x+t) = \sum_{k=1}^{\infty} F_k(x+) = \sum_{k=1}^{\infty} G_k(x)$$

Taking things one step further, define $H_k(x) = G_k(x) - G_k(a)$ and $H(x) = G(x) - G(a)$. Since adding by a constant doesn't change the derivative of a function at all, we still know that $H_k' = F_k'$; $H' = F'$ a.e. on $[a,b]$ Also, since G_k and G are monotone increasing, we know that $G_k(x) \geq G_k(a)$ and $G(x) \geq G(a)$ for all x. Hence, all of our H_k and H are still nonnegative monotone increasing functions on $\mathbb{R}$. And clearly $H(x) = \sum_{k=1}^{\infty} H_k(x)$ for all $x \in \mathbb{R}$.

The significance of this is that if we now prove that $H'(x) = \sum_{k=1}^{\infty} H_k'(x)$ for a.e. $x \in [a,b]$, then we will have also shown that $F'(x) = \sum_{k=1}^{\infty} F_k'(x)$ for a.e. $x \in [a,b]$. But importantly, all our H_k and H are in NBV. After all, they are in BV because they are bounded and monotone increasing. Also, they are all right continuous and $H_k(-\infty) = 0 = H(-\infty)$. It then follows that there are unique finite Borel measures μ_{H_k} and μ_H such that $\mu_{H_k}((-\infty, x]) = H_k(x)$ and $\mu_H((-\infty, x]) = H(x)$.

Now let $\mathrm{d}\mu_{H_k} = \mathrm{d}\lambda_k + f_k \mathrm{d}m$ and $\mathrm{d}\mu_H = \mathrm{d}\lambda + f\mathrm{d}m$ be the Radon-Nikodym representations of μ_{H_k} and μ_H with respect to the Lebesgue measure. Then since μ_H and μ_{H_k} are both finite measures in the separable locally compact Hausdorff space $\mathbb{R}$, we know by theorem 7.8 that μ_{H_k} and μ_H are regular. Also, if we let $E_r(x) = (x, x+r]$ for all $r > 0$ and $x \in \mathbb{R}$, then we know that E_r shrinks nicely to x for all x. Therefore, by the generalized Lebesgue differentiation theorem, we have:

$$H'(x) = \lim_{h \to 0^+} \frac{H(x+h) - H(x)}{h} = \lim_{h \to 0^+} \frac{\mu_H((x, x+h])}{h} = \lim_{h \to 0^+} \frac{\mu_H(E_h(x))}{m(E_h(x))} = f(x) \text{ for } m\text{-a.e. } x.$$

(And similarly we have that $H_k'(x) = f_k(x)$ for m-a.e. x.)

Next note that for any $(a,b) \subseteq \mathbb{R}$, we have that:

$$\mu_H((a,b)) = \lim_{\beta \to b-} \mu_H((a,\beta]) = \lim_{\beta \to b-} (H(\beta) - H(a)) = \lim_{\beta \to b-} \sum_{k=1}^{\infty} (H_k(\beta) - H_k(a)).$$

Once again, $H_k(\beta) - H_k(a) \geq 0$ for all k and our series is bounded from above by $\sum_{k=1}^{\infty} (H_k(b) - H_k(a)) = \mu_{H_k}(a,b] < \infty$. So, by applying the dominated convergence theorem we get that:

$$\sum_{k=1}^{\infty} (H_k(\beta) - H_k(a)) = \sum_{k=1}^{\infty} \lim_{\beta \to b-} (H_k(\beta) - H_k(a)) = \sum_{k=1}^{\infty} \lim_{\beta \to b-} \mu_{H_k}((a,\beta]) = \sum_{k=1}^{\infty} \mu_{H_k}((a,b)).$$

This in turn proves that $\mu_H = \sum_{k=1}^{\infty} \mu_{H_k}$ on all open sets by the countable additivity of measures, and all measurable sets in general by outer regularity. Hence, we know that $\mathrm{d}\lambda + H'\mathrm{d}m = \sum_{k=1}^{\infty} (\mathrm{d}\lambda_k + H_k'\mathrm{d}m) = \sum_{k=1}^{\infty} \mathrm{d}\lambda_k + \sum_{k=1}^{\infty} H_k'\mathrm{d}m$.

Lastly, note that $\sum_{k=1}^{\infty} \lambda_k \perp m$ and $\sum_{k=1}^{\infty} H_k'\mathrm{d}m = \left(\sum_{k=1}^{\infty} H_k'\right) \mathrm{d}m \ll m$. By the Radon-Nikodym theorem, we thus know that $\lambda = \sum_{k=1}^{\infty} \lambda_k$ and $H'\mathrm{d}m = \left(\sum_{k=1}^{\infty} H_k'\right) \mathrm{d}m$ with $H' = \sum_{k=1}^{\infty} H_k'$ m-a.e. ■

Exercise 3.40: Let F denote the Cantor function on $[0,1]$ and set $F(x) = 0$ for $x > 1$ and $F(x) = 1$ for $x > 1$. Let $\{[a_n, b_n]\}_{n\in\mathbb{N}}$ be an enumeration of the closed subintervals of $[0,1]$ with distinct rational endpoints, and let $F_n(x) = F(\frac{x-a_n}{b_n-a_n})$. Then $G = \sum_{n=1}^{\infty} 2^{-n} F_n$ is continuous and strictly increasing on $[0,1]$, and $G' = 0$ a.e.

Since $\frac{x-a_n}{b_n-a_n}$ is continuous and increasing, we know that each F_n is still continuous and monotone increasing. Also, we clearly have that if $x \geq b_n$, then $F_n(x) = 1$. Meanwhile, if $x \leq a_n$, then $F_n(x) = 0$. Thus, it's easy to see that:

- $G(x) = 0$ for $x \leq 0$ and $G(x) = 1$ for $x \geq 1$
- G is monotone increasing
- $\sum_{n=1}^{\infty} 2^{-n} F_n$ converges uniformly to G, thus making G continuous.
- By an easy application of exercise 3.39, $G' = 0$ a.e. since F' being zero almost everywhere implies $(2^{-n}F_n)' = 0$ a.e. for each n.

Reminder, for any x not in the Cantor set, we know either $x \notin [0,1]$ or there is an open interval containing x that was removed to form the Cantor set. In either scenario, we have that f is constant on a neighborhood of x. So, $f'(x) = 0$.

Finally, to show that G is strictly increasing on $[0,1]$, note that for any $x, y \in [0,1]$ with $x < y$, we know there is a closed subinterval $[a_n, b_n]$ with $x < a_n < b_n < y$. In turn, we know that $F_n(x) = 0$ while $F_n(y) = 1$. Then since $F_n(y) \geq F_n(x)$ for all other n, we know that $G(x)$ is strictly less than $G(y)$.

Note: We can also fairly easily see now that $\sum_{n\in\mathbb{Z}}(n + G(x-n))\chi_{[n,n+1]}$ is strictly increasing and continuous everywhere with a derivative equal to zero almost everywhere.

This poses a challenge because in exercise 8.18(b), we're going to need to be able say that a function having a derivative of zero almost everywhere implies that the function is constant. So here is one more lemma.

Lemma: If $f : \mathbb{R} \to \mathbb{C}$ is absolutely continuous on $[a,b]$ and $f' = 0$ a.e. on $[a,b]$, then f is constant on $[a,b]$.

Why: By the fundamental theorem of calculus for Lebesgue integrals, we know that if $x \in [a,b]$, then $f(x) - f(a) = \int_a^x f'(t)\mathrm{d}t = \int_a^x 0\mathrm{d}t = 0$. So, $f(x) = f(a)$.

7/13/2025

Exercise 8.18 (continued):

(b) If the L^2 derivative f' exists, then: $\left[\int |f(x)|^2 \mathrm{d}x\right]^2 \leq 4 \int |xf(x)|^2 \mathrm{d}x \int |f'(x)|^2 \mathrm{d}x$.

To start out, we need to make sure this inequality is well defined. Note that since $f, f' \in L^2$, we know that $\int |f(x)|^2 \mathrm{d}x < \infty$ and $\int |f'(x)|^2 \mathrm{d}x < \infty$. So, to guarentee that this inequality is well defined, we just need to show that if $\int |f'(x)|^2 \mathrm{d}x = 0$, then we will never have that $\int |xf(x)|^2 \mathrm{d}x = \infty$ (thus making the right-hand side $4(\infty \cdot 0)$). Luckily, by exercise 8.9, we know that f having an L^2 derivative means that f is absolutely continuous on every bounded interval. So by the lemma I ended yesterday with, we know that if $\int |f'(x)|^2 \mathrm{d}x = 0$, then f must be constant on every bounded interval since the ordinary derivative of f is zero almost everywhere. This proves that $f = c$ where c is some constant. But since $f \in L^2$, we must have that $\int_{-\infty}^{\infty} |c|^2 \mathrm{d}x < \infty$. The only way this is possible is if $c = 0$. So, $f = 0$ a.e. and we've thus shown that $\int |xf(x)|^2 \mathrm{d}x = 0$ as well.

Next, for any $a < b$ note that $|f|^2$ is absolutely continuous on $[a, b]$. This is because as mentioned before, f is absolutely continuous on $[a, b]$. Then in turn, it is easy to see that $\overline{f}$ is absolutely continuous on $[a, b]$. So, by exercise 3.35, we know that $f\overline{f} = |f|^2$ is absolutely continuous on $[a, b]$. Also $g(x) = x$ is absolutely continuous on $[a, b]$. Thus by exercise 3.35, we know that:

$$\int_a^b (1|f(x)|^2 + x\tfrac{\mathrm{d}}{\mathrm{d}x}|f(x)|^2) = b|f(b)|^2 - a|f(a)|^2$$

Or in other words: $\int_a^b |f(x)|^2 \mathrm{d}x = b|f(b)|^2 - a|f(a)|^2 - \int_a^b x\frac{\mathrm{d}}{\mathrm{d}x}|f(x)|^2 \mathrm{d}x$.

Also, note:

$$\tfrac{\mathrm{d}}{\mathrm{d}x}|f(x)|^2 = \tfrac{\mathrm{d}}{\mathrm{d}x}(f(x)\overline{f(x)}) = f'(x)\overline{f(x)} + f(x)\overline{f'(x)} = 2\mathrm{Re}(f'(x)\overline{f(x)}).$$

Hence, for any $a < b$, we have:

$$\int_a^b |f(x)|^2 \mathrm{d}x = b|f(b)|^2 - a|f(a)|^2 - 2\mathrm{Re}(\int_a^b f'(x)\overline{f(x)}\mathrm{d}x)$$

Now since the inequality we want to prove is trivial if $\int |xf(x)|^2 \mathrm{d}x = \infty$, we can safely assume $\int |xf(x)|^2 \mathrm{d}x < \infty$. This is important because it guarentees that for any $n \in \mathbb{N}$ and $\varepsilon > 0$, we can pick $a_n < -n$ and $b_n > n$ such that $|a_n f(a_n)|^2 < {}^1\!/\!_n$ and $|b_n f(b_n)|^2 < {}^1\!/\!_n$. In turn, this lets us say that $a_n|f(a_n)|^2 \in (-{}^1\!/\!_n, 0]$ and $b_n|f(b_n)|^2 \in [0, {}^1\!/\!_n)$ since $a_n < -1$ and $b_n > 1$.

Now by an application of dominated convergence theorem, we know that:

$$\begin{aligned}\int |f(x)|^2 \mathrm{d}x &= \lim_{n\to\infty} \int_{a_n}^{b_n} |f(x)|^2 \mathrm{d}x \\ &= \lim_{n\to\infty} \left(b_n|f(b_n)|^2 - a_n|f(a_n)|^2 - 2\mathrm{Re}(\int_{a_n}^{b_n} f'(x)\overline{f(x)}\mathrm{d}x)\right) \\ &= -2 \cdot \lim_{n\to\infty} \mathrm{Re}(\int_{a_n}^{b_n} xf'(x)\overline{f(x)}\mathrm{d}x)\end{aligned}$$

Then by the Cauchy-Schwartz inequality (using the fact that $f'\chi_{[a_n,b_n]}, x\overline{f} \in L^2$), we can say that:

$$\begin{aligned}-2 \cdot \lim_{n\to\infty} \mathrm{Re}(\int_{a_n}^{b_n} xf'(x)\overline{f(x)}\mathrm{d}x) &\leq 2 \lim_{n\to\infty} \left|\int_{a_n}^{b_n} xf'(x)\overline{f(x)}\right| \\ &\leq 2 \lim_{n\to\infty} \left(\int_{a_n}^{b_n} |f'(x)|^2 \mathrm{d}x\right)^{1/2} (\int |xf(x)|^2 \mathrm{d}x)^{1/2}\end{aligned}$$

By a final application of dominated convergence theorem using an upper bound of $|f'(x)|^2$, we get that:

$$\lim_{n\to\infty} \left(\int_{a_n}^{b_n} |f'(x)|^2 \mathrm{d}x\right)^{1/2} = (\int |f'(x)|^2 \mathrm{d}x)^{1/2}$$

So, $\int |f(x)|^2 \mathrm{d}x \leq (\int |f'(x)|^2 \mathrm{d}x)^{1/2} (\int |xf(x)|^2 \mathrm{d}x)^{1/2}$. Squaring both sides gives the desired inequality.

(c) (**Heisenberg's Inequality**) For any $b, \beta \in \mathbb{R}$,

$$\int (x-b)^2 |f(x)|^2 \mathrm{d}x \int (\xi - \beta)^2 |\widehat{f}(\xi)|^2 \mathrm{d}\xi \geq \tfrac{\|f\|_2^4}{16\pi^2}$$

To start off, $f = 0$ a.e. if and only if $\widehat{f} = 0$ a.e. It follows that we will never have a $0 \cdot \infty$ situation on the left-hand side, and thus our inequality is well-defined. Also, if either of the two left hand integrals are infinite, then the inequality is trivial. So, we may assume both integrals are finite.

Next note that if we consider $g(x) = f(x+b)$, then I already showed on page 71 that $\widehat{g}(\xi) = e^{2\pi i\xi b}\widehat{f}(\xi)$. In turn, we know that:

- $\int x^2|g(x)|^2\mathrm{d}x = \int x^2|f(x+b)|^2\mathrm{d}x = \int (x-b)^2|f(x)|^2\mathrm{d}x$,
- $\int (\xi-\beta)^2|\widehat{g}(\xi)|^2\mathrm{d}\xi = \int (\xi-\beta)^2|e^{2\pi i\xi b}\widehat{f}(\xi)|^2\mathrm{d}\xi = \int (\xi-\beta)^2|e^{2\pi i\xi b}\widehat{f}(\xi)|^2\mathrm{d}\xi$,
- $\|g\|_2 = \|f\|_2$.

So, by proving our inequality for g when $b = 0$, we've also proven it for f when b is anything.

Going a step further, set $h(x) = e^{-2\pi i\beta x}g(x)$. Then $h^\vee(\xi) = g^\vee(\xi-\beta) = \widehat{g}(\beta-\xi)$. So, we know that $\widehat{h}(\xi) = \widehat{g}(\beta+\xi)$ since $\widehat{g}(\xi) = g^\vee(-\xi)$. In turn:

- $\int x^2|h(x)|^2\mathrm{d}x = \int x^2|e^{-2\pi i\beta x}g(x)|^2\mathrm{d}x = \int x^2|g(x)|^2\mathrm{d}x$,
- $\int \xi^2|\widehat{h}(\xi)|^2\mathrm{d}\xi = \int \xi^2|\widehat{g}(\xi+\beta)|^2\mathrm{d}\xi = \int (\xi-\beta)^2|\widehat{g}(\xi)|^2\mathrm{d}\xi$,
- $\|h\|_2 = \|g\|_2$.

So, by proving our inequality for h when $b = 0$ and $\beta = 0$, we've also proven it for f when b and β are anything. Luckily, proving that for h is easy due to what we've already proven in parts (a) and (b) of this exercise.

Since $\int \xi^2|\widehat{h}(\xi)|^2\mathrm{d}\xi < \infty$, we know from part (a) that h has an L^2 derivative h' which satisfies that:

$$\tfrac{1}{2\pi i\xi}\widehat{h'}(\xi) = \widehat{h}(\xi).$$

In turn, we can rewrite $\int \xi^2|\widehat{h}(\xi)|^2\mathrm{d}\xi = \frac{1}{4\pi^2}\int |\widehat{h'}(\xi)|^2\mathrm{d}\xi$ and the latter is just $\frac{1}{4\pi^2}\int |h'(\xi)|^2\mathrm{d}\xi$ by the Plancharel theorem. Finally, by applying part (b) we get that:

$$\tfrac{1}{4\pi^2}\int x^2|h(x)|^2\mathrm{d}x \int |h'(\xi)|^2\mathrm{d}\xi \geq \tfrac{\|h\|_2^4}{4}\cdot\tfrac{1}{4\pi^2} = \tfrac{\|h\|_2^4}{16\pi^2}. \blacksquare$$

This inequality is the cause of the quantum uncertainty principle. To see why, first note that in quantum mechanics, a property of a particle at a given point in time is modeled as a probability density function whose density at a point x is $|f(x)|^2$ where f is some function in L^2 (importantly this means $\|f\|_2 = 1$ always in this context).

In turn, $\int (x-b)^2|f(x)|^2\mathrm{d}x$ is the formula for the variance of that probability distribution around b. So, that integral evaluates to something small precisely when the probability distribution of the property of the particle has a small standard deviation and b is close to the mean of the distribution.

Next, note that in quantum mechanics, pairs of properties are related to each other by a Fourier transformation. Hence, $|\widehat{f}(\xi)|^2 \mathrm{d}x$ is the probability density function of another property of the particle.

Similarly to before, $\int (x-\beta)^2 |\widehat{f}(x)|^2 \mathrm{d}x$ is the formula for the variance of that probability distribution around β, and that will be small precisely when the probability distribution of the property has a small standard deviation and β is close to the mean of the distribution.

Now finally, $\int (x-b)^2|f(x)|^2\mathrm{d}x \int (x-\beta)^2|\widehat{f}(x)|^2\mathrm{d}x \geq \frac{1}{16\pi^2}$ for all $b, \beta \in \mathbb{R}$ means that it's impossible for both probability distributions to simultaneously have a standard deviation less than $\frac{1}{2\sqrt{\pi}}$, and decreasing one of the standard deviations beyond that value necessarily requires increasing the other. This is the quantum uncertainty principle.

Exercise 8.19: If $f \in L^2(\mathbb{R}^n)$ and the set $S = \{x : f(x) \neq 0\}$ has finite measure, then for any measurable $E \subseteq \mathbb{R}^n$, $\int_E |\widehat{f}|^2 \leq \|f\|_2^2 m(S)m(E)$.

By Minkowski's inequality for integrals, we have:

$$\begin{aligned}\int_E |\widehat{f}|^2 &= \int \chi_E(\xi) |\int f(x) e^{-2\pi i \xi\cdot x}\mathrm{d}x|^2 \mathrm{d}\xi \\ &= \int |\int f(x)e^{-2\pi i\xi\cdot x}\sqrt{\chi_E(\xi)}\mathrm{d}x|^2\mathrm{d}\xi \\ &\leq \left[\int (\int |f(x)e^{-2\pi i \xi\cdot x}\sqrt{\chi_E(\xi)}|^2 \mathrm{d}\xi)^{1/2}\mathrm{d}x\right]^2 = \left[\int |f(x)|(\int_E \mathrm{d}\xi)^{1/2}\mathrm{d}x\right]^2 = m(E)\left(\int |f(x)|\mathrm{d}x\right)^2\end{aligned}$$

Next, by Hölder's inequality we have:

$$\int |f(x)|\mathrm{d}x = \int |\chi_S(x) f(x)|\mathrm{d}x \leq \|\chi_S\|_2 \|f\|_2 = \sqrt{m(S)}\|f\|_2.$$

Thus $\int_E |\widehat{f}|^2 \leq m(E)(\sqrt{m(S)}\|f\|_2)^2 = m(E)m(S)\|f\|_2^2$. ■

This inequality is another cause/statement of the quantum uncertainty principle. This is because to optimize the precision of our measurement of the property associated to the wave $|\widehat{f}|^2$, we'd want to maximize $\int_E |\widehat{f}|^2$ while simultaneously minimizing $m(E)$. But, this inequality says that doing that requires increasing $m(S)$. I.e., it requires us to know less about the property associated to the wave $|f|^2$.

7/15/2025

Welp, I'm currently sick. Anyways, now that I've scanned my paper notes from winter quarter, I'm thinking I want to finally learn vector calculus properly since I never really learned it in math 20E. Also, since I'm crashing the physics 4 sequence, I'm going to eventually need to finally learn Stokes' theorem and Divergence theorem when they cover E&M.

For now, my plan is to sort of follow along with Munkres' Analysis on Manifolds, starting at chapter 5. That said, I want to work with Lebesgue integrals. So, I might go on some tangents and or come up with different proofs for things. Also, I might skip something if I'm bored.

Conventions:

- $\{e_1, \ldots, e_n\}$ will refer to the standard basis on $\mathbb{R}^n$.
- If $f \in C^r(U)$ where $U \subseteq \mathbb{R}^n$ is open and $r \geq 1$, then $\mathrm{D}f$ will refer to the derivative matrix of f with respect to the standard bases. I.e, $\mathrm{D}f$ is the matrix of partial derivatives of f.

Lemma: Let W be a k-dimensional linear subspace of $\mathbb{R}^n$. Then there is an orthogonal (i.e. unitary) linear transformation $h : \mathbb{R}^n \to \mathbb{R}^n$ that carries W onto the subspace $\mathbb{R}^k \times 0^{n-k}$ of $\mathbb{R}^n$.

Proof:
Let $\{v_1, \ldots, v_n\}$ be an orthonormal basis for W such that $\{v_1, \ldots, v_k\}$ is an orthonormal basis for W. Then if g is the linear map whose matrix with respect to the standard basis has columns, $v_1, \ldots, v_n$, we know g is orthogonal and $g(e_i) = v_i$ for all i. Now just set $h = g^{-1}$.

Theorem: Let $k, n \in \mathbb{N}$ with $0 < k \leq n$. There is a unique function V that assigns to each k-tuple: $(x_1, \ldots, x_k)$, of elements in $\mathbb{R}^n$ a nonnegative number such that:

1. If $h : \mathbb{R}^n \to \mathbb{R}^n$ is an orthogonal transformation, then:
$$V(h(x_1), \ldots, h(x_k)) = V(x_1, \ldots, x_k)$$
2. If $y_1, \ldots, y_k$ belong to the subspace $\mathbb{R}^k \times 0^{n-k}$ with $y_i = \left[\begin{smallmatrix} z_i \\ 0 \end{smallmatrix}\right]$ where $z_i \in \mathbb{R}^k$, then $V(y_1, \ldots, y_k) = |\det(z_1, \ldots, z_k)|$.

Specifically, we have, $V(x_1, \ldots, x_k) = \left(\det(X^\mathsf{T} X)\right)^{1/2}$ where X is the $n \times k$ matrix $X = [x_1, \ldots, x_k]$.

(Note: we will typically abreviate $V(x_1, \ldots, x_k)$ as $V(X)$...)

Proof:
Given $X = [x_1, \ldots, x_k]$, define $F(X) = \det(X^\mathsf{T} X)$. Then note:

- If $h : \mathbb{R}^n \to \mathbb{R}^n$ is an orthogonal linear map, then $h(x) = Ax$ where A is an orthogonal matrix. Now $[h(x_1), \ldots, h(x_k)] = [Ax_1, \ldots, Ax_k] = AX$. Thus:
$F(h(X)) = f(AX) = \det{(AX)^\mathsf{T} AX} = \det X^\mathsf{T} A^\mathsf{T} AX = \det X^\mathsf{T} X = F(X)$
- If Z is a $k \times k$ matrix and Y is the $n \times k$ matrix $\left[\begin{smallmatrix} Z \\ 0 \end{smallmatrix}\right]$, then:
$$F(Y) = \det\left(\begin{bmatrix} Z^\mathsf{T} & 0 \end{bmatrix} \begin{bmatrix} Z \\ 0 \end{bmatrix}\right) = \det\left(Z^\mathsf{T} Z\right) = (\det(Z))^2$$

With that, all we need to do is show that F is nonnegative so that we can take the square root of F. Luckily, if $\{x_1, \ldots, x_k\}$ are any k-tuple of vectors in $\mathbb{R}^n$, then we know there is a k-dimensional linear subspace of $\mathbb{R}^n$ containing $x_1, \ldots, x_k$. By our prior lemma, there exists an orthogonal linear map h taking W to $\mathbb{R}^k \to 0^{n-k}$. Next by our first bullet, we know that $F(h(X)) = F(X)$. And finally, by our second bullet we know $F(h(X)) = (\det(h(X)))^2 \geq 0$.

Now just define $V(X) = \sqrt{F(X)}$. This proves the existence part of the theorem.

To prove uniqueness, suppose U also satisfies our two axioms. Then for any $\{x_1, \ldots, x_k\} \subseteq \mathbb{R}^n$, let $h(x) = Ax$ be an orthogonal linear map taking $\{x_1, \ldots, x_k\}$ into $\mathbb{R}^k \times 0^{n-k}$. Thus $U(X) = U(h(X)) = |\det(AX)| = V(h(X)) = V(X)$

Side note: if $\{x_1, \ldots, x_k\}$ are not linearly independent, then $V(X) = 0$.

Also, hopefully it's clear that the significance of V is that we now have a way of defining the k-dimensional volume of a k-dimensional parallelepiped in $\mathbb{R}^n$.

With that, we're ready to start integrating on manifolds. We'll start with the simple case of a manifold parametrized by a single function.

Let $k \leq n$. Let A be open in $\mathbb{R}^k$ and let $\alpha : A \to \mathbb{R}^n$ be an injective C^r map (where $r \geq 1$). The set $Y = \alpha(A)$ together with the map α constitute a parametrized-manifold of dimension k. We denote this parametrized manifold Y_α. For a topology, we equip Y_α with the subspace topology of Y with respect to $\mathbb{R}^n$. That way α is still a continuous map.

Next, we define a Borel measure on Y_α. Given any set $E \in \mathcal{B}_{Y_\alpha}$, we define:

$$V(E) := \int_{\alpha^{-1}(E)} V(\mathrm{D}\alpha)$$

(unfortunately the measure is typically called V even though we already named another function that.)

Note, $V(\mathrm{D}\alpha)$ is Borel measurable because the matrix determinant is a continuous function with respect to all the matrix entries and all the entries of $\mathrm{D}\alpha$ are continuous since $\alpha \in C^1$. It follows that our integral is well-defined.

Since $V(\emptyset) = 0$ and V is clearly countably additive, we know V is a measure. Also, the naturalness of this measure is hopefully clear. After all, you can imagine that we are approximating the k-dimensional volume using a bunch of tiny parallelepipeds.

Then, given any measurable function $f : Y \to \mathbb{C}$ on our manifold, we can integrate via the formula: $\int_{Y_\alpha} f \mathrm{d}V = \int_A (f \circ \alpha) V(\mathrm{D}\alpha) \mathrm{d}m$.

Why: This formula clearly holds for simple functions. Then you can extend that to nonnegative functions via the monotone convergence theorem and then to real and complex functions in the standard way.

Theorem: Let $g : A \to B$ be a diffeomorphism of open sets in $\mathbb{R}^k$ and let $\beta : B \to \mathbb{R}^n$ be an injective C^r map (with $r \geq 1$). If we define $\alpha = \beta \circ g$, then $\alpha : A \to \mathbb{R}^n$ is also an injective C_r map and $\alpha(A) = \beta(B) = Y$. Then, a function $f : Y \to \mathbb{C}$ is integrable over Y_β if and only if it is integrable over Y_α, in which case:

$$\int_{Y_\alpha} f \mathrm{d}V_\alpha = \int_{Y_\beta} f \mathrm{d}V_\beta$$

Proof:
The measurability of f is independent of our parametrization since the topology of Y_α and Y_β was not defined using α or β. Next note that by change of variables:

$$\int_{Y_\beta} f\mathrm{d}V_\beta = \int_B (f\circ\beta)V(\mathrm{D}\beta)\mathrm{d}m = \int_A (f\circ\beta\circ g)(V(\mathrm{D}\beta)\circ g)|\det(\mathrm{D}g)|\mathrm{d}m$$
$$= \int_A (f\circ\alpha)(V(\mathrm{D}\beta)\circ g)|\det(\mathrm{D}g)|\mathrm{d}m$$

Thus, we just need to show that $(V(\mathrm{D}\beta)\circ g)|\det(\mathrm{D}g)| = V(\mathrm{D}\alpha)$. To do that, note by chain rule that $\mathrm{D}\alpha = ((\mathrm{D}\beta)\circ g)\mathrm{D}g$. Therefore:

$$(V(\mathrm{D}\alpha))^2 = \det((((\mathrm{D}\beta)\circ g)\mathrm{D}g)^\mathsf{T}(((\mathrm{D}\beta)\circ g)\mathrm{D}g))$$
$$= (\det(\mathrm{D}g))^2\det(((\mathrm{D}\beta)\circ g)^\mathsf{T}((\mathrm{D}\beta)\circ g)) = (\det(\mathrm{D}g))^2(V((\mathrm{D}\beta)\circ g))^2 \blacksquare$$

Exercise 22.1: Let A be open in $\mathbb{R}^k$, $\alpha : A \to \mathbb{R}^n$ be a C^1 map, and $Y = \alpha(A)$. Suppose $h : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry and let $Z = h(Y)$ and $\beta = h\circ\alpha$. Then Y_α and Z_β have the same volume.

Since h is an isometry, we know h has the form $h(x) = Qx + b$ where Q is an orthogonal matrix and b is a constant vector. Thus by chain rule, we have that $\mathrm{D}\beta = Q\mathrm{D}\alpha$. So:

$$\int_{Z_\beta}\mathrm{d}V = \int_A V(\mathrm{D}\beta) = \int_A V(Q\mathrm{D}\alpha) = \int_A V(\mathrm{D}\alpha) = \int_{Y_\alpha}\mathrm{d}V$$

7/16/2025

Now our previous approach to manifolds is lacking in some respects. For one, it'd be nice if our manifolds were able to include their boundaries. However, requiring the domains of our parametrizations to be strictly open makes that difficult. Also, it'd be nice if we could talk about manifolds that can't be parametrized by a single function (such as the unit sphere S^2 in $\mathbb{R}^3$).

To address the first issue, we extend our notion differentiability. Let S be any subset of $\mathbb{R}^k$ and let $f : S \to \mathbb{R}^n$. We say $f \in C^r(S)$ iff there exists an open set $U \supset S$ and a function $g : U \to \mathbb{R}^n$ in $C^r(U)$ such that $g|_S = f$.

Note that if $f_1, f_2 \in C^r(S)$, then we have that $f_1 + f_2 \in C^r(S)$ and $f_1f_2 \in C^r(S)$. After all, supposing we can extend f_1 and f_2 to the functions g_1 and g_2 defined respectively on the open sets U and V containing S, then $U\cap V$ is still open and contains S, and $g_1+g_2, g_1g_2 \in C^r(U\cap V)$.

Also, if $f_1 \in C^r(S)$ and $f_2 \in C^r(T)$ where $f_1(S) \subseteq T$, then $f_2\circ f_1 \in C^r(S)$. This is because if g_1 extends f_1 to an open set $U \supseteq S$ and g_2 extends f_2 to an open set $V \supseteq T$, then $g_2 \circ g_1$ is a C^r extension of $f_2\circ f_1$ to the open set $g_1^{-1}(V) \supseteq S$.

Next, we show that being in $C^r(S)$ is a local property. However, to do that we first prove a result about partitions of unity.

Recall from Math 240B that a partition of unity on a set $E \subseteq X$ (where X is an LCH space such as $\mathbb{R}^n$) is a collection of functions $\{h_\alpha\}_{\alpha\in A} \in C(X,[0,1])$ such that:

- Each $x \in X$ has a neighborhood on which only finitely many h_α are nonzero,
  Although Munkres only requires this to hold for all $x \in E$.
- $\sum_{\alpha\in A} h_\alpha(x) = 1$ for all $x \in E$.

Also, we say $\{h_\alpha\}_{\alpha\in A}$ is subordinate to the open cover $\mathcal{U}$ of E if for each α there exists $U \in \mathcal{U}$ such that $\mathrm{supp}(h_\alpha) \subseteq U$.

A really cool result Munkres spends all of chapter 16 proving is that for any collection $\mathcal{A}$ of open sets in $\mathbb{R}^n$ whose union is A, there is a partition of unity $\{h_m\}_{m\in\mathbb{N}}$ of A consisting of C_c^∞ functions and such that $\{h_m\}_{m\in\mathbb{N}}$ is subordinate to $\mathcal{A}$.

Lemma 16.2: Let $\mathcal{A}$ be a collection of open sets in $\mathbb{R}^n$ whose union is A. Then there exists a countable collection $\{Q_i\}_{i\in\mathbb{N}}$ of rectangles (i.e. Cartesian products of closed intervals) contained in A such that:

1. The sets $\{Q_i^\circ\}_{i\in\mathbb{N}}$ cover A.
2. Each Q_i is contained in an element of $\mathcal{A}$.
3. Each point of A has a neighborhood that intersects only finitely many of the sets Q_i.

Proof:

Step 1: Dividing A into a nicely structured sequence of compact sets:

Because $\mathbb{R}^n$ is σ-compact, we know A is a countable union of compact sets. Then, by taking larger and larger finite unions of those compact sets we get an increasing sequence of compact sets $\{K_i\}_{i\in\mathbb{N}}$ whose union is A.

Next, to get a nicer sequence, we do more finagling. Let $D_1 = K_1$. Then for $i \geq 1$, define D_{i+1} inductively as follows:

We know there exists a precompact open set V such that $(D_i \cup K_i) \subseteq V \subseteq \overline{V} \subseteq A$. Thus, set $D_{i+1} = \overline{V}$.

Thus, $\{D_i\}_{i\in\mathbb{N}}$ is a sequence of compact sets whose union is A and which satisfies that $D_i \subseteq D_{i+1}^\circ$ for all i. Also, for convenience of notation let $D_i = \emptyset$ if $i \leq 0$.

Finally, set $B_i = D_i - D_{i-1}^\circ$ for all i. Then note that $\{B_i\}_{i\in\mathbb{N}}$ is a sequence of compact sets whose union is A and which satisfies that B_i is disjoint from D_{i-2} for all $i \geq 2$. Consequently, this means that any B_i only intercepts B_{i-1} and B_{i+1} in our sequence. Also, $U_i \coloneqq D_{i+1}^\circ - D_{i-2}$ is an open neighborhood of B_i which intercepts only B_{i-1}, B_i, and B_{i+1}.

Step 2: Making our covering of A:

After fixing i, note that for any $x \in B_i$ we know that there is a set $E \in \mathcal{A}$ such that $x \in E$. So, we can pick a rectangle Q_x such that $x \in Q_x^\circ$ and $Q_x \subseteq E \cap U_i$. Doing this for all $x \in B_i$, we get a collection $\{Q_x\}_{x \in B_i}$ of sets whose interiors give an open covering of B_i. So, because B_i is compact, there is a finite collection $\mathcal{C}_i := \{Q_{x_1}, \ldots, C_{Q_{n_i}}\}$ of rectangles whose interiors cover all of B_i and such that each rectangle is a subset of some element of $\mathcal{A}$ intercepted with U_i.

Repeat this process for all i and let $\mathcal{C} = \bigcup_{i \in \mathbb{N}} \mathcal{C}_i$. Then $\mathcal{C}$ is a countable collection of rectangles such that each $Q \in \mathcal{C}$ is contained in an element of $\mathcal{A}$. Also, for any $x \in A$, we know there is some j with:

$$x \in (B_{j-1} \cup B_j \cup B_{j+1}) - \bigcup_{\substack{i \in \mathbb{N} \\ i \notin \{j-1, j, j+1\}}} B_i$$

In turn, x is in the interior of one the rectangles in $\mathcal{C}_{j-1} \cup \mathcal{C}_j \cup \mathcal{C}_{j+1}$.

Finally, supposing $x \in B_k$, then we know the open neighborhood $U_k \subseteq A$ of x only intercepts U_{k-1} and U_{k+1}, which in turn only intercept B_{k-2} through B_{k+2}. So, U_k is an open neighborhood of x which intercepts at most the rectangles from $\mathcal{C}_{j-2}$ through $\mathcal{C}_{j+2}$. ■

<u>Theorem 16.3:</u> Let $\mathcal{A}$ be a collection of open sets in $\mathbb{R}^n$ whose union is A. Then there is a partition of unity $\{h_m\}_{m \in \mathbb{N}}$ subordinate to $\mathcal{A}$ such that each h_m is in C_c^∞.

Proof:

Construct $\{Q_m\}_{m \in \mathbb{N}}$ like in the previous lemma. Then note that for each Q_m, there is a C_c^∞ function g_m such that $g_m(x) > 0$ if $x \in Q_m^\circ$ and $g_m(x) = 0$ otherwise.

Specifically: If you define $f(x) = e^{1/x}$ when $x > 0$ and $f(x) = 0$ when $x \leq 0$, then $f \in C^\infty(\mathbb{R})$. In turn, $g := f(x)f(1-x) \in C^\infty(\mathbb{R})$ with $g(x) > 0$ when $x \in (0,1)$ and $g(x) = 0$ otherwise.

Finally, if $Q_m = [a_1, b_1] \times \ldots \times [a_n, b_n]$, then define:

$$g_m(x_1, \ldots, x_n) = g(\tfrac{x_1 - a_1}{b_1 - a_1}) \cdots g(\tfrac{x_n - a_n}{b_n - a_n})$$

Thus $g_m \in C^\infty(\mathbb{R}^n)$ with $g_m(x) > 0$ when $x \in Q_m^\circ$ and $g_m(x) = 0$ otherwise.

Having done that, we've now guarenteed that $\{g_m\}_{m \in \mathbb{N}}$ has all the properties we want except that we don't necessarily have that $\sum_{m \in \mathbb{N}} g_m(x) = 1$ for all $x \in A$. To fix that, we normalize our functions.

Let $\lambda(x) := \sum_{m=1}^\infty g_m(x)$. Then for any $x \in A$, we know there is an open neighborhood N_x of x that intercepts only finitely many $\mathrm{supp}(g_m)$. It follows then that $\lambda(x) < \infty$ for all $x \in A$ and that λ is infinitely differentiable for all $x \in A$. Meanwhile, since any $x \in A$ is in the interior of at least one Q_m, we know that $\lambda(x) > 0$ for all $x \in A$.

Now for each m define $h_m(x) = g_m(x)/\lambda(x)$ when $x \in A$ and $h_m(x) = 0$ when $x \notin A$. Then it's clear that $\sum_{m\in\mathbb{N}} h_m(x) = 1$ when $x \in A$. Also, we still have that each $h_m \in C^\infty$. To see that, first note that h_m is infinitely differentiable via quotient rule on A. Also, since Q_m is compact, A^C is closed, and both are disjoint, we know there is some minimum distance δ between the two sets. So for any $x \in A^\mathsf{C}$, we know that h_m is just the zero function while on a ball of radius $\delta/2$ around x. So, all of the derivatives of h_m exist and equal zero at x for any $x \in A^\mathsf{C}$. Finally, since $\mathrm{supp}(h_m) = \mathrm{supp}(g_m)$, we have that h_m satisfies our other requirements.

Now returning to our goal of extending the concept of differentiability, we have the following result:

Lemma 23.1: Let S be a subset of $\mathbb{R}^k$ and let $f : S \to \mathbb{R}^n$. If for each $x \in S$ there is a neighborhood U_x of x and a C^r function $g_x : U_x \to \mathbb{R}^n$ which agrees with f on $U_x \cap S$, then $f \in C^r(S)$.

Proof:

For each $x \in S$, pick a set U_x and a C^r function $g_x : U_x \to \mathbb{R}^n$ as allowed by the hypothesis of the lemma. Then set $\mathcal{A} = \{U_x : x \in S\}$ and call the union of that collection of sets A. Via the prior result, there is a partition of unity $\{\phi_m\}_{m\in\mathbb{N}}$ on A consisting of C_c^∞ functions and which is subordinate to $\mathcal{A}$. In turn, for any $m \in \mathbb{N}$ there exists x_m with $\mathrm{supp}(\phi_m) \subseteq U_{x_m}$. It then follows that $h_m \coloneqq \phi_m g_{x_m} \in C^r(U_m)$ and that h_m vanishes outside a compact subset of U_m, meaning we can extend h_m to being in $C^r(\mathbb{R}^k)$ by setting $h_m = 0$ outside U_{x_m}.

Finally, define $g = \sum_{m=1}^{\infty} h_m$ on A. Then for any $x \in A$, we know that x has a neighborhood on which g is only a sum of finitely many h_m. So, $g \in C^r(A)$. Also, if $x \in S$, then for any m with $\phi_m(x) \neq 0$, $h_m(x) = \phi_m(x)g_m(x) = \phi_m(x)f(x)$. Therefore, for any $x \in A \cap S$:

$$g(x) = \sum_{m\in\mathbb{N}} h_m(x) = f(x) \sum_{m\in\mathbb{N}} \phi_m(x) = f(x). \blacksquare$$

7/18/2025

Let $H^k = \{(x_1, \ldots, x_k) \in \mathbb{R}^k : x_k \geq 0\}$ denote the "upper-Half-space". Also, let $H_+^k = \{(x_1, \ldots, x_k) \in \mathbb{R}^k : x_k > 0\}$ denote the "open upper-Half-space" in $\mathbb{R}^k$.

Lemma 23.2: Let U be open in H^k but not in $\mathbb{R}^k$, and let $\alpha : U \to \mathbb{R}^n$ be in $C^r(U)$. That way, there exists a C^r extension $\beta : U' \to \mathbb{R}^n$ of α defined on an open set U' of $\mathbb{R}^k$. Then for $x \in U$, the derivative $\mathrm{D}\beta(x)$ depends only on the function α and is independent of the extension β. Hence it follows we may denote this derivative by $\mathrm{D}\alpha(x)$ without ambiguity.

Why: We know that $\frac{\partial}{\partial x_k}\beta$ is fully determined by the right-hand limit:

$$\lim_{h\to 0^+} \frac{\beta(x+he_k)-\beta(x)}{h} = \lim_{h\to 0^+} \frac{\alpha(x+he_k)-\alpha(x)}{h}$$

It follows that all first order partial derivatives of β on U are uniquely determined by α. Then proceeding by induction and noting that $\partial^\gamma \beta$ extends $\partial^\gamma \alpha$ from U to U' for all multi-indices γ of degree less than r, we can apply the same reasoning to conclude that all partial derivatives of β with order less than r on U are uniquely determined by α.

With that, we now have the ability to define parametrized manifolds with a boundary by making the domain of our parametrization open with respect to H^k instead of $\mathbb{R}^k$. As for our other issue of wanting to talk about manifolds that can't be parametrized by a single function, we shall deal with that now.

Let $k > 0$. A k-manifold in $\mathbb{R}^n$ of class C^r is a subset M of $\mathbb{R}^n$ having the following property: For each $p \in M$ there is an open set V of M containing p, a set U that is open in either $\mathbb{R}^k$ or H^k, and a continuous bijective map $\alpha : U \to V$ such that:

1. $\alpha \in C^r(U)$
2. $\alpha^{-1} : V \to U$ is continuous
3. $\mathrm{D}\alpha(x)$ has rank k for each $x \in U$.

The map α is called a coordinate patch on M about p.

Also, we call a discrete collection of points in $\mathbb{R}^n$ to be a 0-manifold.

Lemma 23.3: Let M be a manifold in $\mathbb{R}^n$ and $\alpha : U \to V$ be a coordinate patch on M. If U_0 is a subset of U that is open in U, then the restriction of α to U_0 is also a coordinate patch on M.

Proof:
Since α^{-1} is continuous and U_0 is open in U, we know that $V_0 := \alpha(U_0)$ is also open in V and thus also M. Hence, $\alpha|_{U_0}$ is a coordinate patch on M because it carries U_0 onto V_0 in a bijective fashion, and it's a C^r map with a continuous inverse and $\mathrm{D}(\alpha|_{U_0})$ having rank k just because it's the restriction of α which has all of those things.

Exercise 23.3(b): Why is $\alpha : [0, 1) \to \mathbb{R}^2$ defined by $\alpha(t) = (\cos(2\pi t), \sin(2\pi t))$ not a coordinate patch for the unit circle S^1?

In this example, α^{-1} is not continuous at $\alpha(0) = (1, 0)$. One way to see this is that the limit of α^{-1} going one way around the circle towards $(1, 0)$ will be 1, whereas the limit going the other way around the circle will be 0.

It should be noted though that S^1 is still a 1-manifold. It's just that we need to use multiple overlapping coordinate patches that don't individually go all the way around the circle in order to cover it.

In order to prove the next theorem, I actually need to generalize the version of the inverse function theorem that I learned in 140C so that if f is a bijective C^r map with a nonsingular derivative, then I know that $g = f^{-1}$ is also C^r rather than just merely C^1. But to do that, I need to finally learn Cramer's rule.

I'm also realizing right about now that in my original notes where I defined $\det$ (my MITx notes which I just got back from my parents, yay!), while my construction still generalizes to matrices defined on arbitrary scalar fields perfectly well, it does have a slight problem of using the parity of permutations in it's definition. However, in Math 100A we defined the parity of a permutation by taking the determinant of it's matrix representation. So, I might as well deal with that cyclic definition now...

Given a permutation $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$, we can define the function:

$$\mathrm{sgn}(\sigma) = \prod_{i<j} \mathrm{sgn}(\sigma(j) - \sigma(i))$$

Now I claim that sgn is a group homomorphism from S_n to the multiplicative group $\{-1, 1\} \subseteq \mathbb{R}^n$ To see this, suppose $\sigma \in S_n$ and that $\tau_{k_1,k_2} \in S_n$ is the transposition swapping k_1 and k_2. Then it's clear that:

$$\mathrm{sgn}(\tau_{k_1,k_2} \circ \sigma) = -\mathrm{sgn}(\sigma) = \mathrm{sgn}(\tau_{k_1,k_2})\mathrm{sgn}(\sigma).$$

In turn, if $\sigma' \in S_n$ is another arbitrary permutation, then by expressing $\sigma' = \tau_1 \circ \tau_2 \circ \ldots \circ \tau_N$ where all the τ_i are transpositions, we have that:

$$\begin{aligned}
\mathrm{sgn}(\sigma' \circ \sigma) &= \mathrm{sgn}(\tau_1 \circ \tau_2 \circ \ldots \circ \tau_{N-1} \circ \tau_N \circ \sigma) \\
&= \mathrm{sgn}(\tau_1)\mathrm{sgn}(\tau_2 \circ \ldots \circ \tau_{N-1} \circ \tau_N \circ \sigma) \\
&\vdots \\
&= \mathrm{sgn}(\tau_1)\mathrm{sgn}(\tau_2) \ldots \mathrm{sgn}(\tau_{N-1})\mathrm{sgn}(\tau_N)\mathrm{sgn}(\sigma) \\
&= \mathrm{sgn}(\tau_1)\mathrm{sgn}(\tau_2) \ldots \mathrm{sgn}(\tau_{N-1} \circ \tau_N)\mathrm{sgn}(\sigma) \\
&\vdots \\
&= \mathrm{sgn}(\tau_1)\mathrm{sgn}(\tau_2 \circ \cdots \tau_{N-1} \circ \tau_N)\mathrm{sgn}(\sigma) \\
&= \mathrm{sgn}(\tau_1 \circ \tau_2 \circ \cdots \tau_{N-1} \circ \tau_N)\mathrm{sgn}(\sigma) = \mathrm{sgn}(\sigma')\mathrm{sgn}(\sigma)
\end{aligned}$$

Also, it's easily checked that $\mathrm{sgn}(\mathrm{Id}) = 1$. Thus sgn is a group homomorphism. And, since every transposition has a negative sign, we get the following nice interpretation of the sign of a permutation. Specifically: $\mathrm{sgn}(\sigma) = 1$ if σ can only be constructed using an even number of transpositions starting from the identity, and $\mathrm{sgn}(\sigma) = -1$ if σ can only be constructed using an odd number of transpositions starting from the identity.

Next, here are Cramer's rules:

Theorem 2.13: Let $A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$ be an $n \times n$ matrix. Also let:

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ and } c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Then if $Ax = c$, we have:

$$\det(A)x_i = \det\left(\begin{bmatrix} a_1 & \cdots & a_{i-1} & c & a_{i+1} & \cdots & a_n \end{bmatrix}\right)$$

Proof:
Define $C = \begin{bmatrix} e_1 & \cdots & e_{i-1} & x & e_{i+1} & \cdots & e_n \end{bmatrix}$.

Then, we have that $AC = \begin{bmatrix} a_1 & \cdots & a_{i-1} & c & a_{i+1} & \cdots & a_n \end{bmatrix}$. Thus, we know:
$$\det(A)\det(C) = \det\left(\begin{bmatrix} a_1 & \cdots & a_{i-1} & c & a_{i+1} & \cdots & a_n \end{bmatrix}\right).$$

Also note that $\det(C) = x_i(-1)^{i+i}\det(I_{n-1}) = x_i(-1)^{2i} = x_i$. The desired conclusion then follows.

Theorem 2.14: Let A be an $n \times n$ matrix of rank n and let $B = A^{-1} = [b_{i,j}]$. Then letting $A_{j,i}$ denote the matrix which results from removing the jth row and ith column of A, we have that:
$$b_{i,j} = \frac{(-1)^{j+i}\det(A_{j,i})}{\det(A)}$$

Proof:
After fixing j, set $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ equal to the jth column of B.

Now since $AB = I_n$, we know $Ax = e_j$. Therefore, by our last theorem we know:
$$\det(A)x_i = \det\left(\begin{bmatrix} a_1 & \cdots & a_{i-1} & e_j & a_{i+1} & \cdots & a_n \end{bmatrix}\right)$$

The latter determinant equals $(-1)^{j+i}\det(A_{j,i})$. Therefore:
$$x_i = b_{i,j} = \frac{(-1)^{j+i}\det(A_{j,i})}{\det(A)}. \blacksquare$$

Now, here's the generalization of the inverse function theorem, that if $f \in C^r$ and satisfies all the other hypotheses of the inverse function theorem, then we can guarentee that our inverse function $g = f^{-1}$ is also C^r.

Proof:
Jumping to where we ended our proof of the inverse function theorem in math 140C, we had shown that $\mathrm{D}g(y) = (\mathrm{D}f(g(y)))^{-1}$ for all y in some open subset of the image of f. If $g(y) = (g_1(y), \ldots, g_n(y))$, then by our prior theorem we know that:
$$\frac{\partial}{\partial y_j}g_i(y) = \frac{(-1)^{j+i}\det([\mathrm{D}f(g(y))]_{j,i})}{\det(\mathrm{D}f(g(y)))}$$

Now we already know that g is C^1. So, when proceeding by induction for $r > 1$, it suffices to assume g is also C^{r-1}. But then note that since $\det(\mathrm{D}f(g(y))) \neq 0$ for all y, and since all the partial derivatives of f are C^{r-1}, our above expression shows that $\frac{\partial}{\partial y_j}g_i(y)$ is also C^{r-1}. And since this works for all partial derivatives of g, we've proven that g is $C^{(r-1)+1}$. $\blacksquare$

And finally to finish off for tonight...

Theorem 24.1 Let M be a k-manifold in $\mathbb{R}^n$ of class C^r. Also let $a_0 : U_0 \to V_0$ and $a_1 : U_1 \to V_1$ be coordinate patches on M with $W = V_0 \cap V_1$ nonempty, and let $W_i = \alpha_i^{-1}(W)$. Then the map $(\alpha_1^{-1} \circ \alpha_0) : W_0 \to W_1$ is a C^r map with a nonsingular derivative.

(Side note: We often call $\alpha_1^{-1} \circ \alpha_0$ the transition function between the coordinate patches α_0 and α_1.)

Proof:
It suffices to show that if $\alpha : U \to V$ is a coordinate patch on M, then $\alpha^{-1} : V \to \mathbb{R}^k$ is in $C^r(V)$.

Why: If α_0 and α_1^{-1} are both of class C^r, then so is their composite $\alpha_1^{-1} \circ \alpha_0$. By similar reasoning, we also know that $\alpha_0^{-1} \circ \alpha_1 : W_1 \to W_0$ is in $C^r(W_0)$. And since $(\alpha_1^{-1} \circ \alpha_0)$ and $\alpha_0^{-1} \circ \alpha_1$ are inverses of each other, we know by chain rule that for any $x \in W_0$ and $y = (\alpha_1^{-1} \circ \alpha_0)(x)$:

$$\mathrm{D}(\alpha_0^{-1} \circ \alpha_1)(y)\mathrm{D}(\alpha_1^{-1} \circ \alpha_0)(x) = \mathbf{1}$$

The only way this is possible is if $\det(\mathrm{D}(\alpha_1^{-1} \circ \alpha_0)) \neq 0$ for all $x \in W_0$.

Next, to prove that α^{-1} is of class C^r, it suffices to show that it is locally of class C^r. So let p_0 be a point of V and set $x_0 = \alpha^{-1}(p_0)$.

First consider the case U is open in H^k but not in $\mathbb{R}^k$. Then, we can extend α to a C^r map β on an open set U' of $\mathbb{R}^k$. Now $\mathrm{D}\alpha(x_0)$ has rank k. So after some suitable permutation of our standard basis vectors, we can assume the first k rows of the matrix $\mathrm{D}\alpha(x_0)$ are linearly independent. If we then define $\pi : \mathbb{R}^n \to \mathbb{R}^k$ to be the projection from $\mathbb{R}^n$ onto those first k basis vector coordinates, we have that the map $g = \pi \circ \beta$ maps U' into $\mathbb{R}^k$ and $\mathrm{D}g(x_0)$ is non-singular. So by the inverse function theorem, we know g is a C^r diffeomorphism on an open set W of $\mathbb{R}^k$ about x_0 (meaning g is C^r, g has an inverse g^{-1}, and g^{-1} is C^r).

Now we claim $h = g^{-1} \circ \pi$ (which is a C^r map) extends α^{-1} to a neighborhood A of p_0. Firstly note $U_0 := W \cap U$ is open in U. Hence, α^{-1} being continuous implies that $V_0 := \alpha(U_0)$ is open in V. This means there is an open set $A \subseteq \mathbb{R}^n$ such that $A \cap V = V_0$. By intercepting A with $\pi^{-1}(g(W))$ ($= \beta(W)$), we can force A to be contained in the domain of h.

Now $h : A \to \mathbb{R}^k$ is of class C^r, and if $p \in A \cap V = V_0$, then when letting $x = \alpha^{-1}(p)$ we have:

$$h(p) = h(\alpha(x)) = g^{-1}(\pi(\alpha(x))) = g^{-1}(g(x)) = x = \alpha^{-1}(p).$$

As for the case where U is open in $\mathbb{R}^k$, then just set $U' = U$ and $\beta = \alpha$ and the prior reasoning still works. ■

Side note: As a corollary, we now know that if two coordinate patches parametrize the same manifold, then the domains of those two coordinate patches are diffeomorphic. So hopefully that adds to the significance of theorem I wrote at the bottom of page 79.

7/19/2025

Today I shall formalize what I mean by the "boundary" of a manifold and maybe also do some cool exercises. I'll try to get as much done as possible since I'm going to be busy at San Diego pride tomorrow.

Let M be a k-manifold in $\mathbb{R}^n$ and let $p \in M$. If there is a coordinate patch $\alpha : U \to V$ on M about p such that U is open in $\mathbb{R}^k$, we say p is an interior point of M. Otherwise, we say p is a boundary point of M.

We denote the set of boundary points of M as ∂M and call it the boundary of M. Meanwhile, we call $M - \partial M$ the interior of M. Note that these definitions are distinct from the topological definitions of boundaries and interiors.

> Untill I get bored of this and go back to Folland or someone else, I'll be using ∂M to refer to the manifold definition of boundary as opposed to a different definition.

Lemma 24.2: Let M be a k-manifold in $\mathbb{R}^n$ and $\alpha : U \to V$ be a coordinate patch about the point p of M.

(a) If U is open in $\mathbb{R}^k$, then p is an interior point of M.
(b) If U is open in H^k and $p = \alpha(x_0)$ where $x_0 \in H^k_+$, then p is an interior point of M.
(c) If U is open in H^k and $p = \alpha(x_0)$ where $x_0 \in \mathbb{R}^{k-1} \times 0$, then p is a boundary point of M.

Proof:
Parts (a) and (b) are trivial. As for part (c), let $\alpha_0 : U_0 \to V_0$ be the coordinate patch in the hypothesis of the lemma and suppose (for the sake of contradiction) that there is another coordinate patch $\alpha_1 : U_1 \to V_1$ about p with U_1 open in $\mathbb{R}^k$.

Since V_0 and V_1 are open in M, the set $W = V_0 \cap V_1$ is also open in M. Let $W_i = \alpha_i^{-1}(W)$ for $i = 0, 1$. Then W_0 is open in H^k and contains x_0 (which consequently means W_0 isn't open in $\mathbb{R}^k$). Also, W_1 is open in $\mathbb{R}^k$. But now note that by our prior theorem, $\alpha_0^{-1} \circ \alpha_1$ is a C^r map from W_1 to $W_0 \subseteq \mathbb{R}^k$ with a nonsingular derivative matrix. So, by specifically part (A) of the inverse function theorem (as covered in math 140C), we know W_1 maps to an open set in $\mathbb{R}^k$. Yet $\alpha_0^{-1} \circ \alpha_1(W_1) = W_0$ is not open in $\mathbb{R}^k$. Hence, a contradiction. ■

> Side note: Holy fuck I did not realize before now that in math 140C we proved that C^1 functions to $\mathbb{R}^k$ with a nonsingular derivative matrix are open maps.

Note, we trivially have that H^k is a k-manifold of class C^∞ in $\mathbb{R}^k$ (just define the coordinate patch to be the identity map on H^k.) Then, $\partial H^k = \mathbb{R}^{k-1} \times 0$ by the prior lemma.

Theorem 24.3: Let M be a k-manifold in $\mathbb{R}^n$ of class C^r. If ∂M is nonempty, then ∂M is a $k-1$ manifld without boundary in $\mathbb{R}^n$ of class C^r.

Proof:
Let $p \in \partial M$, and then let $\alpha : U \to V$ be a coordinate patch on M about p. Then U is open in H^k and $p = \alpha(x_0)$ for some $x_0 \in \partial H^k$. By the prior lemma, $\alpha(x) \in \partial M$ for all $x \in U \cap \partial H^k$ and $\alpha(x) \notin \partial M$ for all $x \in U - \partial H^k$. Thus, we know that the restriction of α to $U \cap \partial H^k$ is a bijective map onto the open set $V_0 \coloneqq V \cap \partial M$ of ∂M.

Now let U_0 be the open set in $\mathbb{R}^{k-1}$ such that $U_0 \times 0 = U \cap \partial H^k$. Then for any $x \in U_0$, define $\alpha_0(x) = \alpha(x, 0)$. Thus, $\alpha_0 : U_0 \to V_0$ is a coordinate patch on ∂M about p.

- It is C^r because so is α.
- $\mathrm{D}\alpha_0(x)$ has rank $k-1$ for all x since $\mathrm{D}\alpha_0$ just consists of the first $k-1$ columns of $\mathrm{D}\alpha(x, 0)$.
- Finally, α_0^{-1} is continuous because it equals the composition of α^{-1} restricted to the set V_0 followed by the projection of $\mathbb{R}^k$ onto it's first $(k-1)$-coordinates (and both of those functions are continuous).

This proves ∂M is a manifold. Also, this shows that p is an interior point of ∂M. So, ∂M has no boundary.

Theorem 24.4: Let $\mathcal{O}$ be an open set in $\mathbb{R}^n$, and let $f : \mathcal{O} \to \mathbb{R}$ be a C^r map. Also let M be the set of points for which $f(x) = 0$ and N be the set of points for which $f(x) \geq 0$. If $M \neq \emptyset$ and $\mathrm{D}f(x)$ has rank 1 for all x in M, then N is an n-manifold in $\mathbb{R}^n$ and $M = \partial N$.

Consequently, a level set of f is a manifold so long as f has no critical points in that level set.

Proof:
Firstly, suppose $p \in N$ with $f(p) > 0$. Then if α is the identity map on the set $U \coloneqq f^{-1}((0, \infty))$, we have that α is a C^∞ bijective map from the open set U in $\mathbb{R}^n$ to itself such that α has a continuous inverse and a full rank derivative matrix. So, α is a coordinate patch on N.

Meanwhile, suppose $f(p) = 0$. Then since $\mathrm{D}f(p) \neq 0$, at least one partial derivative $\frac{\partial}{\partial x_i} f(p)$ is nonzero. By a sufficient permutation of our standard basis vectors, we can assume $i = n$. So, define $F : \mathcal{O} \to \mathbb{R}^n$ by the equation $F(x) = (x_1, \ldots, x_{n-1}, f(x))$. Thus, F is a C^r map with a nonsingular derivative matrix at p since:

$$\mathrm{D}F = \begin{bmatrix} I_{n-1} & 0 \\ * & \frac{\partial}{\partial x_n} f \end{bmatrix}$$

By the inverse function theorem, we know F is a C^r diffeomorphism from an open neighborhood V of p in $\mathbb{R}^n$ to an open set U of $\mathbb{R}^n$. Furthermore, F carries the open set $V \cap N$ of N onto the open set $U \cap H^n$ of H^n. Therefore, $F^{-1}|_{(U \cap H^n)}$ works as our coordinate patch on N about p.

Finally note that $F(p) \in \partial H^n$. This shows that $M = \partial N$. ■

Corollary 24.5: The n-ball $B^n(a) := \{x : \|x\|_2 \leq a\}$ is a C^∞ n-manifold whose boundary is $S^{n-1} := \{x : \|x\|_2 = a\}$.

Proof:
Consider the function $f(x) = a^2 - (\|x\|_2)^2$.

The next exercise gives us an important tool for constructing manifolds (which makes it kinda shocking that Munkres leaves it as an exercise).

Exercise 24.2 Let $f : \mathbb{R}^{n+k} \to \mathbb{R}^n$ be of class C^r. Let M be the set of all x such that $f(x) = 0$. Assume that M is nonempty and $\mathrm{D}f(x)$ has rank n for all $x \in M$. Then M is a k-manifold in $\mathbb{R}^{n+k}$ without boundary. Furthermore, if N is the set of all x such that $f_1(x) = \cdots = f_{n-1}(x) = 0$ and $f_n(x) \geq 0$, and the matrix $\partial(f_1, \ldots, f_{n-1})/\partial x$ has rank $n-1$ at each point of N, then N is a $k+1$ manifold and $M = \partial N$.

Lemma: Let $m \leq n$ and suppose $f : \mathbb{R}^{n+k} \to \mathbb{R}^m$ is a C^r map such that the matrix $\mathrm{D}f$ has rank m at the point p. Then there are open sets $U, V \subseteq \mathbb{R}^{n+k}$ with $p \in V$, as well as a C^r diffeomorphism $G : U \to V$ with rank $n+k$ satisfying that $f \circ G(x) = \pi_m(x)$ where π_m is a projection from $\mathbb{R}^{n+k}$ to m of it's coordinates (and by applying a suitable permutation of our basis vectors, we can assume that those coordinates are the first m coordinates).

Proof:
Since $\mathrm{D}f$ has rank m at p, we know that the derivative matrix has m linearly independent columns at p, and by a permutation of our standard bases, we can assume those m columns are the first m columns. Therefore, it makes sense to adopt the notation of writing $x = (x^{(1)}, x^{(2)})$ in $\mathbb{R}^{n+k}$ where $x^{(1)} \in \mathbb{R}^m$ and $x^{(2)} \in \mathbb{R}^{n+k-m}$. Also, it makes sense to define the projection $\pi_m(x) = x^{(1)}$.

Now define the function $F : \mathbb{R}^{n+k} \to \mathbb{R}^{n+k}$ by:

$$F(x) = (f_1(x), \ldots, f_m(x), x^{(2)}).$$

Then F is a C^r map with the derivative matrix:

$$\mathrm{D}F = \begin{bmatrix} \partial(f_1, \ldots, f_m)/\partial x^{(1)} & \partial(f_1, \ldots, f_m)/\partial x^{(2)} \\ 0 & I_{n+k-m} \end{bmatrix}$$

Since $\mathrm{D}F$ is full rank at p, by the inverse function theorem there are open sets $U, V \subseteq \mathbb{R}^{n+k}$ with $p \in V$ such that F is a C^r diffeomorphism from V to U with full rank. Taking $G := F^{-1}$, we get:

$$f \circ G(x) = (\pi_m \circ F) \circ G(x) = \pi_m \circ (F \circ F^{-1})(x) = \pi_m(x) = x^{(1)}$$

Given $f = (f_1, \ldots, f_{n-1}, f_n)$, we'll denote $\tilde{f} = (f_1, \ldots, f_{n-1})$. Now, the rest of the exercise just involves applying the prior lemma three times.

1. Suppose $p \in N - M$ (meaning $f_n(p) > 0$ and $\tilde{f}(p) = 0$). Then using our prior lemma, there are open sets $U, V \subseteq \mathbb{R}^{n+k}$ with $p \in V$ as well as a C^r diffeomorphism $\tilde{G} : U \to V$ with full rank such that $\tilde{f} \circ \tilde{G}(x) = \pi_{n-1}(x)$. Using the continuity of f_n, we can assume $f_n(x) > 0$ for all $x \in V$ by making V sufficiently small.

   Now note that $V \cap N$ is an open subset of N containing p such that:
   $$\begin{aligned} V \cap N &= \tilde{G}(U) \cap \tilde{f}^{-1}(\{0\}) \cap \{x : f_n(x) > 0\} \\ &= \tilde{G}(U) \cap (\pi_{n-1} \circ \tilde{G}^{-1})^{-1}(\{0\}) \cap \{x : f_n(x) > 0\} \\ &= \tilde{G}(U) \cap \tilde{G}(\pi_{n-1}^{-1}(\{0\})) \cap \{x : f_n(x) > 0\} \\ &= \tilde{G}\left(U \cap (0^{n-1} \times \mathbb{R}^{k+1})\right) \cap \{x : f_n(x) > 0\} \\ &= \tilde{G}\left(U \cap (0^{n-1} \times \mathbb{R}^{k+1}) \cap \tilde{G}^{-1}(\{x : f_n(x) > 0\})\right) \end{aligned}$$

   Then since $U \cap \tilde{G}^{-1}(\{x : f_n(x) > 0\})$ is open in $\mathbb{R}^{n+k}$, we can deduce that there is an open set $A \subseteq \mathbb{R}^{k+1}$ with $0^{n-1} \times A = U \cap (0^{n-1} \times \mathbb{R}^{k+1}) \cap \tilde{G}^{-1}(\{x : f_n(x) > 0\})$. And by defining $\alpha(x_1, \ldots, x_{k+1}) = \tilde{G}(0^{n-1}, x_1, \ldots, x_{k+1})$, we have that α is a bijective C^r map from the open set A of $\mathbb{R}^{k+1}$ to $V \cap N$ with rank $k+1$ and a continuous inverse. Hence, α is
   a coordinate patch on N about p.

2. Suppose $p \in M$ and only assume the part of the exercise statement that comes before the word "furthermore". Then let $U, V \subseteq \mathbb{R}^{n+k}$ be open sets with $p \in V$, and let $G : U \to V$ be a C^r diffeomorphism with full rank satisfying that $f \circ G(x) = \pi_n(x)$. Then $V \cap M$ is an open subset of M containing p such that:
   $$\begin{aligned} V \cap M &= G(U) \cap f^{-1}(\{0\}) \\ &= G(U) \cap (\pi_n \circ G^{-1})^{-1}(\{0\}) = G(U) \cap G(\pi_n^{-1}(\{0\})) \\ &= G(U \cap (0^n \times \mathbb{R}^k)) \end{aligned}$$

   Now we know there is some open set $A \subseteq \mathbb{R}^k$ such that $0^n \times A = U \cap (0^n \times \mathbb{R}^k)$. Therefore, by defining $\alpha(x_1, \ldots, x_k) = \tilde{G}(0^n, x_1, \ldots, x_k)$, we have that α is a bijective C^r map from the open set A of $\mathbb{R}^{k+1}$ to $V \cap M$ with rank k and a continuous inverse. Hence, α is a coordinate patch on M about p.

3. Finally, suppose $p \in M$ and this time assume the entire hypothesis of the exercise. Also let $G : U \to V$ be as in the prior part. Now:
   $$\begin{aligned} V \cap N = G(U) \cap f^{-1}(\{0\}) &= G(U) \cap G(\pi_n^{-1}(0^{n-1} \times [0, \infty))) \\ &= G(U \cap (0^{n-1} \times [0, \infty) \times \mathbb{R}^k)) \end{aligned}$$

   Now if τ is the function permuting the first and $(k+1)$th basis vectors, then we know there is some open set $A \subseteq H^{k+1}$ such that $0^{n-1} \times \tau(A) = U \cap (0^{n-1} \times [0, \infty) \times \mathbb{R}^k)$. So, define $\alpha(x_1, \ldots, x_{k+1}) = G(0^{n-1}, x_{k+1}, x_1, \ldots, x_k)$. Then α is a bijective C^r map from the open set A of H^{k+1} to $V \cap N$ with rank $k+1$ and a continuous inverse. Hence, α is a coordinate patch on N about p.

Also, if $x \in U$ satisfies that $\alpha(x) = p$, then since $f \circ G = \pi_n$, we know that $f(G(x)) = f(p) = 0$. So $x \in \partial H^{k+1}$. This proves p is on the boundary of N. ■

Note from 8/28/2025: How the fuck did I just realize this is just implicit function theorem.

7/28/2025

Here's a fun application of the prior exercise.

Exercise 24.6: Let $\mathcal{O}(n)$ denote the set of all orthogonal n by n matrices, considered as a subspace of $\mathbb{R}^N$ where $N = n^2$. Show that $\mathcal{O}(n)$ is a compact $\binom{n}{2}$-manifold of class C^∞ in $\mathbb{R}^N$ without boundary.

Firstly, define $F : \mathbb{R}^N \to \mathbb{R}^N$ by $F(\mathbf{A}) = \mathbf{A}\mathbf{A}^T - I_n$. Also define $f : \mathbb{R}^N \to \mathbb{R}^{\frac{1}{2}n(n+1)}$ such that $f(\mathbf{A})$ is the vector containing just the elements of $F(\mathbf{A})$ on or above the main diagonal. Then it is fairly clear that $F(\mathbf{A}) = 0$ if and only if $f(\mathbf{A}) = 0$ if and only if $\mathbf{A} \in \mathcal{O}(n)$.

Next note that if $\mathbf{A} = [x_{i,j}]$, then the (i,j)th. column of the derivative matrix $\mathrm{D}F$ is:

$$\begin{aligned}
\frac{\partial}{\partial x_{i,j}} F(\mathbf{A}) &= \left(\frac{\partial}{\partial x_{i,j}}\mathbf{A}\right)\mathbf{A}^T + \mathbf{A}\left(\frac{\partial}{\partial x_{i,j}}\mathbf{A}\right)^T \\
&= \begin{bmatrix} & \mathbf{0}^{(i-1)\times n} & \\ x_{1,j} & \cdots & x_{n,j} \\ & \mathbf{0}^{(n-i)\times n} & \end{bmatrix} + \begin{bmatrix} & \mathbf{0}^{(i-1)\times n} & \\ x_{1,j} & \cdots & x_{n,j} \\ & \mathbf{0}^{(n-i)\times n} & \end{bmatrix}^T \\
&= \begin{bmatrix} & & x_{1,j} & & \\ & & \vdots & & \\ x_{1,j} & \cdots & 2x_{i,j} & \cdots & x_{n,j} \\ & & \vdots & & \\ & & x_{n,j} & & \end{bmatrix}
\end{aligned}$$

Meanwhile, the (i,j)th. column of $\mathrm{D}f$ is just the vector with the elements of $\mathrm{D}F$ on or above the diagonal. From this, it is clear that f and F are C^∞ maps.

Taking a different view, if $k \leq \ell$, then the (k,ℓ)th. row of $\mathrm{D}f$ is:

$$\left(\begin{bmatrix} x_{\ell,1} & \cdots & x_{\ell,n} \end{bmatrix}(k\text{th. row})\right) + \left(\begin{bmatrix} x_{k,1} & \cdots & x_{k,n} \end{bmatrix}(\ell\text{th. row})\right)$$

Now let $v_{k,\ell}$ denote the (k,ℓ)th. row of $\mathrm{D}f$; let r_i denote the ith. row of $\mathbf{A}$; and let $\delta_{i,j}(x,y)$ equal 1 if $(x,y) = (i,j)$ and equal 0 otherwise. Then supposing $k_1 \leq \ell_1$ and $k_2 \leq \ell_2$, we have that:

$$v_{k_1,\ell_1} \cdot v_{k_2,\ell_2} = \delta_{k_1,k_2}(r_{\ell_1} \cdot r_{\ell_2}) + \delta_{k_1,\ell_2}(r_{\ell_1} \cdot r_{k_2}) + \delta_{\ell_1,k_2}(r_{k_1} \cdot r_{\ell_2}) + \delta_{\ell_1,\ell_2}(r_{k_1} \cdot r_{k_2})$$

If $\mathbf{A}$ is orthogonal, then this simplifies to:

$$\begin{aligned} v_{k_1,\ell_1} \cdot v_{k_2,\ell_2} &= \delta_{k_1,k_2}\delta_{\ell_1,\ell_2} + \delta_{k_1,\ell_2}\delta_{\ell_1,k_2} + \delta_{\ell_1,k_2}\delta_{k_1,\ell_2} + \delta_{\ell_1,\ell_2}\delta_{k_1,k_2} \\ &= 2\delta_{k_1,k_2}\delta_{\ell_1,\ell_2} + 2\delta_{k_1,\ell_2}\delta_{\ell_1,k_2} \end{aligned}$$

There are two cases where this dot product will be nonzero. The first case is if $k_1 = k_2$ and $\ell_1 = \ell_2$. Meanwhile, the second case is if $k_1 = \ell_2$ and $\ell_1 = k_2$. However, since we are requiring that $k_1 \leq \ell_1$ and $k_2 \leq \ell_2$, the second case actually implies the first case.

This proves that the rows of $\mathrm{D}f$ actually form an orthogonal set of vectors in $\mathbb{R}^N$. Hence, f has full rank on $\mathcal{O}(n)$.

It now follows from the previous exercise that $\mathcal{O}(n)$ is a manifold without boundary in $\mathbb{R}^N$. It's dimension will be $n^2 - \frac{(n+1)n}{2} = \frac{n^2}{2} - \frac{n}{2} = \binom{n}{2}$. Meanwhile, to see that the manifold is compact, note firstly that it is bounded by the set $\{x \in \mathbb{R}^N : \|x\|_\infty \leq 1\}$. Also, the points in the manifold are given by the set $f^{-1}(\{0\})$, and that set is closed since f continuous and $\{0\}$. ■

7/29/2025

I'm gonna finish the current chapter of Munkres and then switch to a different book to learn about differential forms. For today, my agenda is to define integration of scalar-valued functions on general manifolds in $\mathbb{R}^n$.

Let M be a k-manifold of class C^r in $\mathbb{R}^n$ (with $r \geq 1$ and $k \leq n$). Then recall that we already defined integration on M if M is parametrized by a single coordinate patch from an open set of $\mathbb{R}^k$. Hopefully, it's also clear to see that our previous definition works if our coordinate patch is from an open set of H^k. Also, by theorem 24.1 plus the theorem at the bottom of page 79 of this journal, we now know that our definition of the integral is independent of the parametrization we use.

Note from 7/31/2025: Actually I haven't yet showed that the parametrization is independent if the domain of that parametrization is open in H^k but not $\mathbb{R}^k$.

In general though, M probably can't be parametrized by just one coordinate patch. So, we instead bodge our definition using a partition of unity.

<u>Lemma:</u> There is a countable set $\{\alpha_n : U_n \to V_n\}_{n\in\mathbb{N}}$ of coordinate patches on M such that $\bigcup_{n\in\mathbb{N}} V_n = M$.

Proof:

1. Note that topologically speaking, M is a second countable LCH space.

   The fact that M is second countable and Hausdorff is just a consequence of the fact that $\mathbb{R}^n$ is both of those things and M is equipped with the subspace topology.

To show that M is locally compact, note that if $p \in M$, then there is a homeomorphism α from some open set U in $\mathbb{R}^k$ or H^k to an open set $V \subseteq M$ containing p. Then given the $x \in U$ satisfying that $\alpha(x) = p$, there is a compact set $K \subseteq U$ with $x \in K$.

If U is open in $\mathbb{R}^k$, then it's obvious that K exists. Meanwhile, if U is open in H^k, then consider picking a U' which is open in $\mathbb{R}^k$ and satisfies that $U' \cap H^k = U$. Then we know there is a compact set K' such that $x \in K' \subseteq U'$. And since H^k is closed, we can set $K = K' \cap H^k$ and know K is compact.

Next, by Urysohn's lemma, there is a precompact open set V such that $K \subseteq V \subseteq \overline{V} \subseteq U$. Hence, we have that $\alpha(\overline{V})$ is a compact subset of M containing p. Also $\alpha(V)$ is an open subset of $\alpha(\overline{V})$ which contains p. Hence, $\alpha(\overline{V})$ is a compact neighborhood of p.

2. In turn, we know that M is σ-compact. From that hopefully it is obvious how we can get a countable covering of coordinate patches over M. ■

<u>Lemma:</u> Let $\{\alpha_a : U_a \to V_a\}_{a \in A}$ be a collection of coordinate patches on M such that $\bigcup_{a \in A} V_a = M$. Then there exists a countable collection of C^∞ functions $\{\phi_i : \mathbb{R}^n \to \mathbb{R}\}_{i \in \mathbb{N}}$ satisfying that:

- $\phi_i(x) \geq 0$ for all $x \in \mathbb{R}^n$ and $i \in \mathbb{N}$.
- Each $p \in M$ has a neighborhood in M on which only finitely many ϕ_i are nonzero.
- $\sum_{i \in \mathbb{N}} \phi_i(p) = 1$ for all $p \in M$.
- For each $i \in \mathbb{N}$, there is some $a \in A$ such that $\mathrm{supp}(\phi_i) \cap M \subseteq V_a$.

In the future I will just refer to $\{\phi_i\}_{i \in \mathbb{N}}$ as being a partition of unity on M subordinate to our collection of coordinate patches.

Proof:
For each coordinate patch α_a, choose an open set $V_a' \subseteq \mathbb{R}^n$ such that $V_a' \cap M = V_a$. Then by theorem 16.3, we get our desired partition which is subordinate to $\{V_a'\}_{a \in A}$.

<u>Definition:</u> Let $(\phi_i)_{i \in \mathbb{N}}$ be a partition of unity on M subordinate to the collection of coordinate patches $\{\alpha_i : U_i \to V_i\}_{i \in \mathbb{N}}$ which cover M. Without loss of generality, suppose $\mathrm{supp}(\phi_i) \cap M \subseteq V_i$. Then, we define a Borel measure on M by:

$$V(E) := \sum_{i=1}^{\infty} \int_{E \cap V_i} \phi_i \mathrm{d}V_{\alpha_i} = \sum_{i=1}^{\infty} \int_{\alpha_i^{-1}(E \cap V_i)} (\phi_i \circ \alpha_i) V(\mathrm{D}\alpha_i) \mathrm{d}m$$

From here it's pretty obvious that $V(\emptyset) = 0$ and that V is countably additive. Also, similarly to before we can then deduce that:

$$\int_M f \mathrm{d}V = \sum_{i=1}^{\infty} \int_{V_i} f \phi_i \mathrm{d}V_{\alpha_i} = \sum_{i=1}^{\infty} \int_{U_i} (f\phi_i \circ \alpha_i) V(\mathrm{D}\alpha_i) \mathrm{d}m$$

Now our first challenge is to show that this definition is independent of our choice of coordinate patches and partition of unity.

Let $(\psi_i)_{i\in\mathbb{N}}$ be another partition of unity on M subordinate to another collection of coordinate patches $\{\beta_i : U_i' \to V_i'\}_{i\in\mathbb{N}}$ which cover M (and like before suppose $\mathrm{supp}(\psi_i) \cap M \subseteq V_i'$).

Now importantly, by our prior results about integration on manifolds parametrized by single coordinate patches, we know that $\int_{V_i \cap V_j'} f \mathrm{d}V_{\alpha_i} = \int_{V_i \cap V_j'} f \mathrm{d}V_{\beta_j}$ for all integrable f on $V_i \cap V_j'$. Therefore, if $E \in \mathcal{B}_M$, we have that:

$$\begin{aligned}
\sum_{i=1}^{\infty} \int_{V_i} \phi_i \chi_E \mathrm{d}V_{\alpha_i} &= \sum_{i=1}^{\infty} \int_{V_i} \phi_i \chi_E \left(\sum_{j=1}^{\infty} \psi_j \chi_{V_j'} \right) \mathrm{d}V_{\alpha_i} \\
&= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \int_{V_i \cap V_j'} \phi_i \psi_j \chi_E \mathrm{d}V_{\alpha_i} \\
&= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \int_{V_i \cap V_j'} \phi_i \psi_j \chi_E \mathrm{d}V_{\beta_j} \\
&= \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} \int_{V_i \cap V_j'} \phi_i \psi_j \chi_E \mathrm{d}V_{\beta_j} \\
&= \sum_{j=1}^{\infty} \int_{V_j'} \psi_j \chi_E \left(\sum_{i=1}^{\infty} \phi_i \chi_{V_i} \right) \mathrm{d}V_{\beta_j} = \sum_{j=1}^{\infty} \int_{V_j'} \psi_j \chi_E \mathrm{d}V_{\beta_j}
\end{aligned}$$

Side note: we can swap the order of the sums in the second to last line by applying Fubini-Tonelli's theorem.

This shows that our definition of the measure of E is independent of our patches and partition of unity.

One other thing to note is that if M can be parametrized by a single coordinate patch, then this result shows that our definition for parametrized manifolds is compatible with this one. After all, just take $\phi_i = \delta_{1,i}$ (where δ is the Kronecker delta function), and let α_i be the same coordinate patch parametrizing all of our manifold for all i.

7/31/2025

Ok. So in the previous journal entry, I was sorta following along with Munkres while only slightly modifying his thoerems. However, after some thought, I actually want to take a completely different approach to defining a measure on manifolds which is more in line with math 240A. So, I'm going to completely depart from Munkres and do a bunch of stuff entirely on my own (i.e. not based on any people's books).

Let M be a k-manifold of class C^r in $\mathbb{R}^n$ (with $r \geq 1$ and $k \leq n$). Then recall from the last journal entry that M is an LCH second countable topological space. Hence, it follows that M is σ-compact, and from there it follows easily that given any arbitrary collection $\{\alpha_a : U_a \to V_a\}_{a \in A}$ of coordinate patches covering $E \subseteq M$, there is a countable subcover over E.

Also, recall that if $\alpha : U \to V$ is a coordinate patch on M, then we can use the formula $\int_{\alpha^{-1}(E)} V(\mathrm{D}\alpha)\mathrm{d}m$ to calculate the "surface-volume" of any Borel set $E \subseteq V$.

Note that if $\alpha_1 : U_1 \to V_1$ and $\alpha_2 : U_2 \to V_2$ are both coordinate patches on M, $E \subseteq V_1 \cap V_2$ is Borel, and U_1, U_2 are open in $\mathbb{R}^k$, then as we already noted, by theorem 24.1 plus the theorem at the bottom of page 79 of this journal:

$$\int_{\alpha_1^{-1}(E)} V(\mathrm{D}\alpha_1)\mathrm{d}m = \int_{\alpha_2^{-1}(E)} V(\mathrm{D}\alpha_2)\mathrm{d}m.$$

That said, before continuing on I want to show that this equivalence still holds if U_1 or U_2 is open in H^k but not $\mathbb{R}^k$.

To start off, by restricting α_1 and α_2 to their preimages of $V_1 \cap V_2$, we can without loss of generality assume that $V_1 = V_2$. This is important because it makes it so that $\alpha_2^{-1} \circ \alpha_1(U_1) = U_2$ and $\alpha_1^{-1} \circ \alpha_2(U_2) = U_1$.

Now, its impossible for U_1 to be open in $\mathbb{R}^k$ but not U_2, or vice versa. After all, if U_1 is open in $\mathbb{R}^k$, we must have that the transition function $\alpha_2^{-1} \circ \alpha_1 : U_1 \to U_2$ maps U_1 to an open set in $\mathbb{R}^k$. Analogous reasoning using the transition function $\alpha_1^{-1} \circ \alpha_2$ works if U_2 is open in $\mathbb{R}^k$.

Now suppose both U_1 and U_2 are open only in H^k. Then, we can easily see that $W_i := U_i - \partial H^k$ is the interior of U_i in $\mathbb{R}^k$ and that $m(U_i - W_i) = 0$ for both i. Also, since the transition functions map open sets of $\mathbb{R}^k$ to open sets of $\mathbb{R}^k$, we have that $\alpha_2^{-1} \circ \alpha_1(W_1) \subseteq W_2$ and $\alpha_1^{-1} \circ \alpha_2(W_2) \subseteq W_1$. This is enough to say that the transition functions restricted to W_1 and W_2 are a diffeomorphism.

Now finally, by applying the lemma at the bottom of page 79, we have for all functions f which are integrable over V:

$$\begin{aligned}\int_{U_1}(f \circ \alpha_1)V(\mathrm{D}\alpha_1)\mathrm{d}m &= 0 + \int_{W_1}(f \circ \alpha_1)V(\mathrm{D}\alpha_1)\mathrm{d}m \\ &= 0 + \int_{W_2}(f \circ \alpha_2)V(\mathrm{D}\alpha_2)\mathrm{d}m = \int_{U_2}(f \circ \alpha_2)V(\mathrm{D}\alpha_2)\mathrm{d}m\end{aligned}$$

Set $f = \chi_E$ and we are done.

We take the following steps to define a measure on M:

(1) Defining an algebra or even a ring of sets is too much to ask for right now. But, we can at least define a collection of sets $\mathcal{A}$ satisfying that if $A \in \mathcal{A}$ and $E \subseteq A$ is a Borel subset of M, then $E \in \mathcal{A}$. Specifically, let $\mathcal{A}$ be the collection of Borel sets $A \subseteq M$ for which there exists a coordinate patch $\alpha : U \to V$ with $A \subseteq V$.

(2) Next, we define a "premeasure" μ_0 on $\mathcal{A}$. Specifically, for each $A \in \mathcal{A}$, define $\mu_0(A) = \int_{\alpha^{-1}(A)} V(\mathrm{D}\alpha)\mathrm{d}m$ where $\alpha : U \to V$ is some coordinate patch with $A \subseteq V$.

Importantly, even though μ_0 isn't a proper premeasure according to Folland's definition since $\mathcal{A}$ isn't actually an algebra, it is still the case that μ_0 and $\mathcal{A}$ are structured enough that the following is easily seen to hold:

- $\mu_0(\emptyset) = 0$
- If $A, B \in \mathcal{A}$ satisfy $A \subseteq B$. then $\mu_0(A) \leq \mu_0(B)$.
- If $(A_j)_{j\in\mathbb{N}}$ is a sequence of disjoint sets in $\mathcal{A}$ with $\bigcup_{j\in\mathbb{N}} A_j \in \mathcal{A}$, then:

$$\mu_0(\bigcup_{j\in\mathbb{N}} A_j) = \sum_{j=1}^{\infty} \mu_0(A_j).$$

(3) Now, we use μ_0 to define an outer measure on $\mathcal{A}$. For any $E \subseteq M$, define:

$$\mu^*(E) = \inf\left\{\sum_{i=1}^{\infty} \mu_0(A_i) : A_i \in \mathcal{A} \text{ for all } i \text{ and } E \subseteq \bigcup_{i\in\mathbb{N}} A_i\right\}$$

Since any subset of M can be covered by countably many coordinate patches, we know that $\mu^*(E)$ is well defined. The rest of the proof that μ^* is an outer measure is then identical to proposition 1.10 of Folland (which is at the top of page 21 of my Latex math 240a notes).

(4) Lemma: $\mu^*|_{\mathcal{A}} = \mu_0$ and every set in $\mathcal{A}$ is μ^* measurable.

Proof:
To prove the first claim, suppose $E \in \mathcal{A}$ and $(A_m)_{m\in\mathbb{N}}$ is a sequence of sets in $\mathcal{A}$ covering E. It's trivial that $\mu^*(E) \leq \mu_0(E)$. Meanwhile let $B_1 = E \cap A_1$ and $B_m = E \cap A_m - \bigcup_{j=1}^{m-1} A_j$. Then the B_m are each disjoint Borel subsets of $E \in \mathcal{A}$ whose union is all of E. Hence all the B_m are in $\mathcal{A}$ and we have:

$$\mu_0(E) = \mu_0(\bigcup_{m\in\mathbb{N}} B_m) = \sum_{m\in\mathbb{N}} \mu_0(B_m) \leq \sum_{m\in\mathbb{N}} \mu_0(A_m)$$

This shows that $\mu_0(E) \leq \mu^*(E)$, thus proving the first claim.

To show the second claim, suppose $A \in \mathcal{A}$, $E \subseteq X$, and $\varepsilon > 0$. Then there exists a sequence $(B_j)_{j\in\mathbb{N}}$ of sets in $\mathcal{A}$ such that $E \subseteq \bigcup_{j=1}^{\infty} B_j$ and $\sum_{j=1}^{\infty} \mu_0(B_j) \leq \mu^*(E) + \varepsilon$. Importantly, $(B_j \cap A)_{j\in\mathbb{N}}$ and $(B_j - A)_{j\in\mathbb{N}}$ are both sequences of sets in $\mathcal{A}$ covering $E \cap A$ and $E - A$ respectively. Also, $\mu_0(B_j) = \mu_0(B_j \cap A) + \mu_0(B_j - A)$ since $B_j \cap A$ and $B_j - A$ are disjoint sets in $\mathcal{A}$ whose union B_j is also in $\mathcal{A}$. Therefore, we have that:

$$\begin{aligned}\mu^*(E) + \varepsilon &\geq \sum_{j=1}^{\infty} \mu_0(B_j) \\ &= \sum_{j=1}^{\infty} \mu_0(B_j \cap A) + \sum_{j=1}^{\infty} \mu_0(B_j - A) \geq \mu^*(E \cap A) + \mu^*(E - A)\end{aligned}$$

Taking $\varepsilon \to 0$, we have that $\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E - A)$ for all $E \subseteq M$. Thus A is μ^* measurable. ■

(5) We now know by Carathéodory's theorem that if $\mathcal{N}$ is the σ-algebra of μ^*-measurable sets, then $\mu := \mu^*|_{\mathcal{N}}$ is a complete measure on $\mathcal{N}$. Furthermore, $\mathcal{A} \subseteq \mathcal{N}$ with $\mu|_{\mathcal{A}} = \mu_0$.

(6) Lemma: If $\mathcal{B}_M$ is the collection of Borel sets on our manifold M, then $\mathcal{B}_M \subseteq \mathcal{N}$. Hence, we can restrict μ to be a Borel measure if we wanted.

Proof:
If $E \in \mathcal{B}_m$, then let $(A_j)_{j\in\mathbb{N}}$ be a countable covering of E consisting of sets in $\mathcal{A}$. Then $E \cap A_j \in \mathcal{A} \subseteq \mathcal{N}$ for all j and $\bigcup_{j\in\mathbb{N}}(E \cap A_j) = E$. So $E \in \mathcal{N}$.

(7) The next thing we want to do now is show that μ is σ-finite. One reason we want to do this is so that we can apply theorems such as Fubini-Tonelli and Radon-Nikodym. Another reason is so that (as I'll show in the next step) μ is guarenteed to be the unique measure on $(M, \mathcal{N})$ which preserves our definition of the measure of a manifold parametrized by a single coordinate patch. Hence, this construction agrees with what I was doing back when I was loosely following Munkres.

To start, we will show that every point $p \in M$ has a neighborhood $A \in \mathcal{N}$ with finite measure.

Let $\alpha : U \to V$ be a coordinate patch on M about p. Then, given $x \in U$ satisfying that $\alpha(x) = p$, let $K \subseteq U$ be a compact neighborhood of x. Since α is a homeomorphism, it is clear that $A := \alpha(K)$ is a compact set containing p in its interior. Hence, $A \in \mathcal{B}_M \subseteq \mathcal{N}$, A is a neighborhood of p and:

$$\mu(A) = \int_K V(\mathrm{D}\alpha)\mathrm{d}m$$

Now $V(\mathrm{D}\alpha)$ is continuous. So by the extreme value theorem, there exists some $C \geq 0$ such that:

$$\int_K V(\mathrm{D}\alpha)\mathrm{d}m \leq \int_K C\mathrm{d}m = Cm(K)$$

And since $m(K) < \infty$, we've shown that $\mu(A) < \infty$.

Now since M is σ-compact, if we pick a set A_p for each $p \in M$ using the reasoning above, then there is a countable subcovering of the A_p over M. This proves that M is σ-finite. ■

(8) Lemma: If $\nu : \mathcal{N} \to [0, \infty]$ is another measure on $(M, \mathcal{N})$ satisfying that $\nu|_{\mathcal{A}} = \mu_0$, then $\nu = \mu$.

The proof of this is almost identical to that of theorem 1.14 in Folland (bottom of page 24 on my Latex math 240a notes).

The one difference is that if $(A_j)_{j\in\mathbb{N}} \subseteq A$ and $A = \bigcup_{j\in\mathbb{N}} A_j$, then $\nu(A) = \mu(A)$ because $A_m - \bigcup_{j=1}^{m-1} A_j \in \mathcal{A}$ for all m and hence:

$$\nu(A) = \sum_{m=1}^{\infty} \nu(A_m - \textstyle\bigcup_{j=1}^{m-1} A_j) = \displaystyle\sum_{m=1}^{\infty} \mu(A_m - \textstyle\bigcup_{j=1}^{m-1} A_j) = \mu(A). \blacksquare$$

Why did I take this pivot? The reason is that now all of Munkres theorems from chapter 25 of his book are obvious including the final theorem about how one would in practice calculate an integral on M. Also, I got to show that this construction is universal in a sense.

I'm now going to switch over to following Guillemin's Differential Forms since I was recommended this book by another tutor. Once again, I'm not going to perfectly follow the book. But I will be using the book as a loose guide.

Tensors:
Let V be an n-dimensional vector space over a field F. Let V^k be the set of all k-tuples of elements of V. Then a function $T : V^K \to F$ is said to be linear in its ith. variable if for all $u, v_1, \ldots, v_k \in V$ and $a, b \in F$, we have that:

$$\begin{aligned} &T(v_1, \ldots, v_{i-1}, av_i + bu, v_{i+1}, \ldots, v_k) \\ &\qquad = aT(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_k) + bT(v_1, \ldots, v_{i-1}, u, v_{i+1}, \ldots, v_k). \end{aligned}$$

If T is linear in all it's variables, we say T is k-linear, and that T is a k-tensor.

Given $k \geq 1$, we shall denote $\mathcal{L}^k(V)$ to be the set of all k-tensors from V. Also, we shall denote $\mathcal{L}^0(V) \coloneqq F$. Note that we can add together and scale k-linear maps to get more k-linear maps. Thus, $\mathcal{L}^k(V)$ is a vector space.

Side note: $\mathcal{L}^1(V)$ is also called the dual space of V and denoted V^*.

Sorry but my apartment mate sent me down a rabbithole that led me to writing a proof on the following pages. I'll come back to tensors right after I'm done writing the following...

8/5/2024

So for context, my apartment mate was reading through a paper by Joel Spencer (who is famous for using probabilistic methods to study combinatorics and graph theory). This paper, which is titled Balancing Unit Vectors, gives a pretty neat existence proof using probability that given any finite collection $\{u_1, \ldots, u_m\}$ of vectors in $\mathbb{R}^n$ with $\|u_i\|_2 \leq 1$ for all i, there exists coefficients $\varepsilon_1, \ldots, \varepsilon_m \in \{-1, +1\}$ such that: $\|\varepsilon_1 u_1 + \ldots + \varepsilon_m u_m\|_2 \leq \sqrt{n}$. However, my apartment mate was having trouble with the following part of the paper and thus asked for my help with it:

> The theorem quickly follows. A linear algebra argument yields $\alpha_1, \ldots, \alpha_m$ satisfying $\alpha_1 u_1 + \cdots + \alpha_m u_m = 0$ such that all $|\alpha_i| \leqslant 1$ and $\alpha_i = \pm 1$ for all but at most n i's. Reordering vectors for convenience we have

Our issue was figuring out what linear algebra argument Spencer was fucking using. After thinking about it for three days while I was grading tests, I finally came up with a proof:

Theorem: Let $u_1, \ldots, u_m$ be vectors in $\mathbb{R}^n$. Then there are constants $a_1, \ldots, a_m$ satisfying $a_1 u_1 + \ldots + a_m u_m = 0$ such that all $|a_i| \leq 1$ and $a_i = \pm 1$ for all but at most n i's.

Proof:

We'll proceed by an inductive argument. For our base case, let $u_{m+1} := 0$. That way $\sum_{i=1}^{m} 0 u_i = u_{m+1}$. Next, suppose that for $n + 1 \leq k \leq m$, we've shown that there are constants $b_1, \ldots, b_k \in [-1, 1]$ and $\varepsilon_{k+1}, \ldots, \varepsilon_{m+1} \in \{-1, +1\}$ satisfying that:

$$b_1 u_1 + \ldots + b_k u_k = \varepsilon_{k+1} u_{k+1} + \ldots + \varepsilon_{m+1} u_{m+1}.$$

If we consider the matrix $U := [u_1 \quad \cdots \quad u_k]$, then letting $b = (b_1, \ldots, b_k)$ we have that any $a = (a_1, \ldots, a_k) \in b + \ker(U)$ will satisfy that:

$$a_1 u_1 + \ldots + a_k u_k = \varepsilon_{k+1} u_{k+1} + \ldots + \varepsilon_{m+1} u_{m+1}.$$

Since $k > n$, we know that $\ker(U)$ is nontrivial and hence unbounded. At the same time, $\|b\|_\infty \leq 1$. Hence, by the connectedness of $b + \ker(U)$ and the continuity of the ∞-norm, we know there is some $a \in b + \ker(U)$ with $\|a\|_\infty = 1$. After reordering our u_i, this is the same as saying that there exists $a_1, \ldots, a_{k-1} \in [-1, 1]$ and $\varepsilon_k = \pm 1$ such that:

$$a_1 u_1 + \ldots + a_{k-1} u_{k-1} + \varepsilon_k u_k = \varepsilon_{k+1} u_{k+1} + \ldots + \varepsilon_{m+1} u_{m+1}.$$

Subtract both sides by $\varepsilon_k u_k$ to complete the induction step.

After induction, we will eventually get constants $a_1, \ldots, a_n \in [-1, 1]$ and $\varepsilon_{n+1}, \ldots, \varepsilon_{m+1}$ equal to ± 1 such that:

$$a_1 u_1 + \ldots + a_n u_n = \varepsilon_{n+1} u_{n+1} + \ldots + \varepsilon_m u_m + \varepsilon_{m+1} u_{m+1}.$$

Move everything over to one side of the equation and forget about the u_{m+1} and we have proven what we wanted.

8/6/2025

Now I'm going to go back to studying tensors.

A multi-index of length k is a k-tuple $I = (i_1, \ldots, i_k)$ of integers. We say I is a multi-index of n if each i is between 1 and n. Now let $u_1, \ldots, u_n$ is a basis of V. For $T \in \mathcal{L}^k(V)$, write $T_I \coloneqq T(u_{i_1}, \ldots, u_{i_k})$ for every multi-index I of n of length k.

Proposition 1.3.7: The T_I uniquely determine T.

Proof:
When $k = 1$, T is just a linear map and we've already proven this for linear maps. For $k > 1$, we proceed by induction. For each i, define $T_i \in L^{k-1}(V)$ by:

$$(v_1, \ldots, v_{k-1}) \mapsto T(v_1, \ldots, v_{k-1}, u_i).$$

Then for $v = c_1 u_1 + \ldots + c_n u_n$, we have:

$$T(v_1, \ldots, v_{k-1}, v) = \sum_{i=1}^{n} c_i T_i(v_1, \ldots, v_{k-1})$$

Also, by induction each T_i is uniquely determined by the coefficients T_I where I is a multi-index of n of length k with a final index equal to i.

Side note: We can see that if $C = (c_{i,j}) \in \mathcal{M}_{n \times k}(F)$ is the matrix satisfying that $v_j = \sum_{i=1}^{n} c_{i,j} u_i$ for each j, then:

$$T(v_1, \ldots, v_k) = \sum_{I=(i_1,\ldots,i_k)} \left(\prod_{j=1}^{k} c_{i_j,j} \right) T_I$$

Given two tensors: $T_1 \in \mathcal{L}^k(V)$ and $T_2 \in \mathcal{L}^\ell(V)$, we define the tensor product of T_1 and T_2 as:

$$(T_1 \otimes T_2)(v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+\ell}) = T_1(v_1, \ldots, v_k) T_2(v_{k+1}, \ldots, v_{k+\ell})$$

Note that:

- $\otimes$ is associative.
- If T_1 or T_2 is a 0-tensor, then $\otimes$ is just scalar multiplication.
- If $\lambda \in F$, $T_1 \in \mathcal{L}^k(V)$, and $T_2 \in \mathcal{L}^\ell(V)$, then $\lambda(T_1 \otimes T_2) = (\lambda T_1) \otimes T_2 = T_1 \otimes (\lambda T_2)$.
- If $T_1, T_2 \in \mathcal{L}^k(V)$ and $T_3 \in \mathcal{L}^\ell(V)$, then $(T_1 + T_2) \otimes T_3 = (T_1 \otimes T_3) + (T_2 \otimes T_3)$. Also $T_3 \otimes (T_1 + T_2) = (T_3 \otimes T_1) + (T_3 \otimes T_2)$

Suppose $T \in \mathcal{L}^k(V)$ and there exists $\ell_1, \ldots, \ell_k \in V^* = \mathcal{L}^1(V)$ such that $T = \ell_1 \otimes \cdots \otimes \ell_k$. Then we say T is decomposable.

If $u_1, \ldots, u_n$ is a basis for V, then we can define the dual basis: $u_1^*, \ldots, u_n^*$ for V^* such that if $v = \sum_{i=1}^{n} c_i u_i$, then $u_j^*(v) = c_j$.

To prove that this is a basis, first note that if $f \in V^*$ and $\lambda_i = f(u_i)$, then we can easily see that for any $v = \sum_{i=1}^{n} c_i u_i$, then:

$$f(v) = \sum_{i=1}^n c_i f(u_i) = \sum_{i=1}^n \lambda_i c_i = \sum_{i=1}^n \lambda_i u_i^*(v).$$

It follow that $u_1^*, \ldots, u_n^*$ span all of V^*. Also, consider if $f = \sum_{i=1}^n \lambda_i u_i^* = 0$ where all the $\lambda_i \in F$. Then we know that $0 = f(u_j) = \lambda_j$ for all j. This shows that $u_1^*, \ldots, u_n^*$ are linearly independent.

Tangent: if V, W are vector fields over F and $A : V \to W$ is a linear map, then we define the transpose $A^\dagger : W^* \to V^*$ of A by $f \mapsto A^\dagger(f) = f \circ A$.

Claim 1.2.15: Suppose $e_1, \ldots, e_m$ is a basis of V and $u_1, \ldots, u_n$ is a basis of W. Then if $A = (a_{i,j})$ is the the matrix of A with respect to the given bases, we have that the matrix of $A^\dagger$ with respect to the dual bases of V and W is given by $(a_{j,i})$.

Proof:

Suppose $(c_{j,i})$ is the matrix representation of $A^\dagger$. Then:

$$A^\dagger(u_i^*)(e_j) = u_i^*(A(e_j)) = u_i^*(\textstyle\sum_{k=1}^n a_{k,j} u_k) = a_{i,j}$$

Simultaneously:

$$A^\dagger(u_i^*)(e_j) = \textstyle\sum_{k=1}^m c_{k,i} e_k^*(e_j) = c_{j,i}$$

Let $u_1, \ldots, u_n$ be a basis of V and $u_1^*, \ldots, u_n^*$ be the corresponding dual basis of V^*. Then for every multi-index $I = (i_1, \ldots, i_k)$ of n of length k, define:

$$u_I^* = u_{i_1}^* \otimes \cdots \otimes u_{i_k}^*.$$

Note that if $J = (j_1, \ldots, j_k)$ is another multi-index of n of length k, then $u_I^*(u_{j_1}, \ldots, u_{j_k}) = \delta_{I,J}$ where δ is the Kronecker delta function.

Theorem 1.3.13: The k-tensors u_I^* form a basis for $\mathcal{L}^k(V)$.

Proof:

Suppose $T \in \mathcal{L}^k(V)$. Then if $T' := \sum_I T_I u_I^*$, we have that $T'_J = T_J$ for all multi-indices J of n of length k. Therefore, since T' and T are uniquely determined by the same T_I, this proves that $T = T'$. So, T is in the span of the u_I^*. This proves that the u_I^* span all of $\mathcal{L}^k(V)$.

Next suppose $T' = \sum_I C_I u_I^* = 0$ where each $C_I \in F$. Then if J is a multi-index of n of length k,

$$0 = T'(u_{j_1}, \ldots, u_{j_k}) = \sum_I C_I u_I^*(u_{j_1}, \ldots, u_{j_k}) = C_J$$

This shows that the u_I^* are linearly independent. ■

Corollary 1.3.15: If V is an n-dimensional vector space, then $\mathcal{L}^k(V)$ is an n^k-dimensional vector space.

If V, W are vector fields over F, $A : V \to W$ is a linear map, and $T \in \mathcal{L}^k(W)$, we define $A^\dagger T(v_1, \ldots, v_k) \coloneqq T(Av_1, \ldots, Av_k)$.

We call $A^\dagger T$ the pullback of T by the map A. Note that this is just a generalization of taking the transpose of A.

Proposition 1.3.18: If we denote $A^\dagger : \mathcal{L}^k(W) \to \mathcal{L}^k(V)$ to be the map $T \mapsto A^\dagger T$, then $A^\dagger$ is a linear map.

This should be pretty obvious...

Furthermore, $A^\dagger(T_1 \otimes T_2) = (A^\dagger T_1) \otimes (A^\dagger T_2)$.

Proof:

Suppose $T_1 \in \mathcal{L}^k(V)$ and $T_2 \in \mathcal{L}^\ell(V)$. Then:

$$\begin{aligned} &A^\dagger(T_1 \otimes T_2)(v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+\ell}) \\ &\qquad = (T_1 \otimes T_2)(Av_1, \ldots, Av_k, Av_{k+1}, \ldots, Av_{k+\ell}) \\ &\qquad = T_1(Av_1, \ldots, Av_k) T_2(Av_{k+1}, \ldots, Av_{k+\ell}) \\ &\qquad = A^\dagger T_1(v_1, \ldots, v_k) A^\dagger T_2(v_{k+1}, \ldots, v_{k+\ell}) \\ &\qquad = (A^\dagger T_1 \otimes A^\dagger T_2)(v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+\ell}). \blacksquare \end{aligned}$$

As a corollary to the above fact, we know that pullbacks map decomposable tensors to decomposable tensors.

Finally, suppose $B : U \to V$ is another linear map. Then $(AB)^\dagger T = B^\dagger(A^\dagger T)$ for all $T \in \mathcal{L}^k(W)$. In other words, $(AB)^\dagger = B^\dagger A^\dagger$.

Proof:

$$\begin{aligned} B^\dagger(A^\dagger T)(v_1, \ldots, v_k) &= A^\dagger T(Bv_1, \ldots, Bv_k) \\ &= T(ABv_1, \ldots, ABv_k) = (AB)^\dagger T(v_1, \ldots, v_k). \blacksquare \end{aligned}$$

Alternating k-Tensors:

Let V be an n-dimensional vector space over a field F with characteristic $\neq 2$ and S_k be the symmetric group over $\{1, \ldots, k\}$. For $\sigma \in S_k$ and $T \in \mathcal{L}^k(V)$, we define $T^\sigma \in \mathcal{L}^k(V)$ by:

$$T^\sigma(v_1, \ldots, v_k) = T(v_{\sigma^{-1}(1)}, \ldots, v_{\sigma^{-1}(k)})$$

Proposition 1.4.14:

(a) If $T = \ell_1 \otimes \cdots \otimes \ell_k$ where each $\ell_i \in V^*$, then $T^\sigma = \ell_{\sigma(1)} \otimes \cdots \otimes \ell_{\sigma(k)}$.

Proof:

$$\begin{aligned} T^\sigma(v_1, \ldots, v_k) = T(v_{\sigma^{-1}(1)}, \ldots, v_{\sigma^{-1}(k)}) &= \ell_1(v_{\sigma^{-1}(1)}) \cdots \ell_k(v_{\sigma^{-1}(k)}) \\ &= \ell_{\sigma(1)}(v_1) \cdots \ell_{\sigma(k)}(v_k) \\ &= (\ell_{\sigma(1)} \otimes \cdots \otimes \ell_{\sigma(k)})(v_1, \ldots, v_k) \end{aligned}$$

(b) If $\sigma \in S_k$, the function $T \mapsto T^\sigma$ is a linear map from $\mathcal{L}^k(V)$ to $\mathcal{L}^k(V)$.

This should be obvious. Also note that this map is invertible via the function $T \mapsto T^{(\sigma^{-1})}$.

(c) If $\sigma, \tau \in S_k$, then $(T^\sigma)^\tau = T^{\sigma\tau}$.

Proof:
Let $u_i \coloneqq v_{\tau^{-1}(i)}$ for all i. Then:

$$\begin{aligned}(T^\sigma)^\tau(v_1, \ldots, v_k) &= T^\sigma(v_{\tau^{-1}(1)}, \ldots, v_{\tau^{-1}(k)}) \\ &= T^\sigma(u_1, \ldots, u_k) = T(u_{\sigma^{-1}(1)}, \ldots, u_{\sigma^{-1}(k)}) \\ &= T(v_{\tau^{-1}(\sigma^{-1}(1))}, \ldots, v_{\tau^{-1}(\sigma^{-1}(k))}) \\ &= T(v_{(\sigma\tau)^{-1}(1)}, \ldots, v_{(\sigma\tau)^{-1}(k)}) \\ &= T^{\sigma\tau}(v_1, \ldots, v_k)\end{aligned}$$

Let V be a vector space and $k \geq 1$ be an integer. Then $T \in \mathcal{L}^k(V)$ is alternating if $T^\sigma = \operatorname{sgn}(\sigma)T$ for all $\sigma \in S_k$. We denote $\mathcal{A}^k(V)$ as the set of all alternating k-tensors on V.

Note:

- If $c_1, c_2 \in F$ and $T_1, T_2 \in \mathcal{A}^k(V)$, then since $T \mapsto T^\sigma$ is a linear map, we have for all $\sigma \in S_k$ that:

$$\begin{aligned}(c_1T_1 + c_2T_2)^\sigma &= c_1T_1^\sigma + c_2T_2^\sigma \\ &= c_1\operatorname{sgn}(\sigma)T_1 + c_2\operatorname{sgn}(\sigma)T_2 = \operatorname{sgn}(\sigma)(c_1T_1 + c_2T_2)\end{aligned}$$

  This proves that $\mathcal{A}^k(V)$ is a subspace of $\mathcal{L}^k(V)$.

- We shall define $\mathcal{A}^0(V) \coloneqq \mathcal{L}^0(V) = F$.

Given an integer $k > 0$, and a tensor $T \in \mathcal{L}^k(V)$, let:

$$\operatorname{Alt}(T) \coloneqq \sum_{\tau \in S_k} \operatorname{sgn}(\tau)T^\tau.$$

Then the alternation operation has the following properties:

Proposition 1.4.17:

(a) Given any $T \in \mathcal{L}^k(V)$ and $\sigma \in S_k$ (where $k > 0$), we have that $(\operatorname{Alt}(T))^\sigma = \operatorname{sgn}(\sigma)\operatorname{Alt}(T)$. I.e., $\operatorname{Alt}(T)$ is an alternating tensor.

Proof:
By proposition 1.4.14 plus the fact that $(\operatorname{sgn}(\sigma))^2 = 1$, we have that:

$$\begin{aligned}(\operatorname{Alt}(T))^\sigma &= \Big(\sum_{\tau \in S_k} \operatorname{sgn}(\tau)T^\tau\Big)^\sigma \\ &= 1 \cdot \sum_{\tau \in S_k} \operatorname{sgn}(\tau)T^{\tau\sigma} = (\operatorname{sgn}(\sigma))^2 \sum_{\tau \in S_k} \operatorname{sgn}(\tau)T^{\tau\sigma} \\ &= \operatorname{sgn}(\sigma) \sum_{\tau \in S_k} \operatorname{sgn}(\tau\sigma)T^{\tau\sigma} = \operatorname{sgn}(\sigma) \sum_{\tau' \in S_k} \operatorname{sgn}(\tau')T^{\tau'} \\ &= \operatorname{sgn}(\sigma)\operatorname{Alt}(T)\end{aligned}$$

(b) If $T \in \mathcal{A}^k(V)$, the $\mathrm{Alt}(T) = k!T$.

Proof:
Since $T^\tau = \mathrm{sgn}(\tau)T$ for all $\tau \in S_k$, we know:

$$\mathrm{Alt}(T) = \sum_{\tau \in S_k} \mathrm{sgn}(\tau)T^\tau = \sum_{\tau \in S_k} (\mathrm{sgn}(\tau))^2 T = \sum_{\tau \in S_k} (1)T = |S_k|T = k!T$$

(c) $\mathrm{Alt}(T^\sigma) = (\mathrm{Alt}(T))^\sigma$.

Proof:
By similar reasoning to in part (a), we have that:

$$\begin{aligned}\mathrm{Alt}(T^\sigma) = 1 \cdot \sum_{\tau \in S_k} \mathrm{sgn}(\tau)T^{\sigma\tau} &= (\mathrm{sgn}(\sigma))^2 \sum_{\tau \in S_k} \mathrm{sgn}(\tau)T^{\sigma\tau} \\ &= \mathrm{sgn}(\sigma) \sum_{\tau \in S_k} \mathrm{sgn}(\sigma\tau)T^{\sigma\tau} \\ &= \mathrm{sgn}(\sigma) \sum_{\tau' \in S_k} \mathrm{sgn}(\tau')T^{\tau'} = \mathrm{sgn}(\sigma)\mathrm{Alt}(T)\end{aligned}$$

And since $\mathrm{sgn}(\sigma)\mathrm{Alt}(T) = (\mathrm{Alt}(T))^\sigma$ by part (a), we know $\mathrm{Alt}(T^\sigma) = (\mathrm{Alt}(T))^\sigma$.

(d) The map $\mathrm{Alt} : \mathcal{L}^k(V) \to \mathcal{L}^k(V)$ defined by $T \mapsto \mathrm{Alt}(T)$ is a linear map. (Also it's onto if F has characteristic 0 or $> k$.)

Proof:
Alt is a linear map because it is a linear combination of a bunch of linear maps.
The onto property follows from part (b).

8/7/2025

If $I = (i_1, \ldots, i_k)$ is a multi-index of n of length k, then we write:

- I is repeating if $i_s = i_r$ for some $s \neq r$.
- I is increasing if $i_1 < i_2 < \ldots < i_k$.
- Given $\sigma \in S_k$, we define $I^\sigma = (i_{\sigma(1)}, \ldots, i_{\sigma(k)})$.

Note that if I is not repeating, then there is a unique permutation $\sigma \in S_k$ such that I^σ is increasing.

Let $u_1, \ldots, u_n$ be a basis of the vector space V over a field F of characteristic $\neq 2$, and let $u_1^*, \ldots, u_n^*$ be the corresponding dual basis. Now given the multi-index $I = (i_1, \ldots, i_k)$, set $u_I^* = u_{i_1}^* \otimes \cdots \otimes u_{i_k}^*$. Next define $\Psi_I = \mathrm{Alt}(u_I^*)$.

Proposition 1.4.20: Let $I = (i_1, \ldots, i_k)$ and $J = (j_1, \ldots, j_k)$ be multi-indices.

(a) $\Psi_{I^\sigma} = \mathrm{sgn}(\sigma)\Psi_I$.

To start off, by the last proposition:

$$\mathrm{sgn}(\sigma)\Psi_I = \mathrm{sgn}(\sigma)\mathrm{Alt}(u_I^*) = (\mathrm{Alt}(u_I^*))^\sigma = \mathrm{Alt}((u_I^*)^\sigma).$$

Next, set $\ell_j = u_{i_j}^*$ for $1 \le j \le k$. Then by proposition 1.4.14, we have:

$$(u_I^*)^\sigma = (\ell_1 \otimes \cdots \otimes \ell_k)^\sigma = \ell_{\sigma(1)} \otimes \cdots \otimes \ell_{\sigma(k)} = u_{i_{\sigma(1)}}^* \otimes \cdots \otimes u_{i_{\sigma(k)}}^* = u_{I^\sigma}^*.$$

Thus $\operatorname{sgn}(\sigma)\Psi_I = \operatorname{Alt}((u_I^*)^\sigma) = \operatorname{Alt}(u_{I^\sigma}^*) = \Psi_{I^\sigma}$.

(b) If I is repeating, $\Psi_I = 0$.

If I is repeating, then there exists $r \neq s$ such that $i_r = i_s$. Then in turn, if $\tau_{r,s} \in S_k$ is the transposition of r and s, then $\operatorname{sgn}(\tau_{r,s}) = -1$ and $I^{\tau_{r,s}} = I$. Then by part (a), we have:

$$\Psi_I = \Psi_{I^{\tau_{r,s}}} = \operatorname{sgn}(\tau_{r,s})\Psi_I = -\Psi_I$$

This is only possible if $\Psi_I(v_1, \ldots, v_k) = 0$ for all $v_1, \ldots, v_k \in V$. Hence Ψ_I is the zero map.

(c) If I and J are strictly increasing, then $\Psi_I(u_{j_1}, \ldots, u_{j_k}) = \delta_{I,J}$ where δ is the Kronecker delta function.

To start off, note that:

$$\begin{aligned}
\Psi_I(u_{j_1}, \ldots, u_{k_k}) &= \sum_{\tau \in S_k} \operatorname{sgn}(\tau)(u_I^*)^\tau(u_{j_1}, \ldots, u_{j_k}) \\
&= \sum_{\tau \in S_k} \operatorname{sgn}(\tau)(u_{i_1}^* \otimes \cdots \otimes u_{i_k}^*)^\tau(u_{j_1}, \ldots, u_{j_k}) \\
&= \sum_{\tau \in S_k} \operatorname{sgn}(\tau)(u_{i_{\tau(1)}}^* \otimes \cdots \otimes u_{i_{\tau(k)}}^*)(u_{j_1}, \ldots, u_{j_k}) \\
&= \sum_{\tau \in S_k} \operatorname{sgn}(\tau)u_{i_{\tau(1)}}^*(u_{j_1}) \cdots u_{i_{\tau(k)}}^*(u_{j_k})
\end{aligned}$$

Now it's clear that $u_{i_{\tau(1)}}^*(u_{j_1}) \cdots u_{i_{\tau(k)}}^*(u_{j_k}) = \delta_{I^\tau,J}$. Also, since both J and I are strictly increasing and also since there is only one permutation such that I^σ is strictly increasing for any nonrepeating I, we know that $\delta_{I^\tau,J} = 1$ iff $\tau = \operatorname{Id}$ and $I = J$. And in that case $\operatorname{sgn}(\tau) = 1$. Hence:

$$\sum_{\tau \in S_k} \operatorname{sgn}(\tau)u_{i_{\tau(1)}}^*(u_{j_1}) \cdots u_{i_{\tau(k)}}^*(u_{j_k}) = \delta_{I,J}.$$

Proposition 1.4.24: Suppose F has characteristic 0 or greater than k. Then $\{\Psi_J : J \text{ is increasing}\}$ is a basis for $\mathcal{A}^k(V)$.

Proof:
Suppose $T \in \mathcal{A}^k(V)$. By theorem 1.3.13, we know there exists $a_I \in F$ such that $T = \sum_I a_I u_I^*$. However, we also know that $k!T = \operatorname{Alt}(T)$. Since Alt is a linear map, we thus know that:

$$T = \frac{1}{k!}\operatorname{Alt}(T) = \sum_I \frac{a_I}{k!}\operatorname{Alt}(u_I^*) = \sum_I \frac{a_I}{k!}\Psi_I$$

If I is repeating, then $\frac{a_I}{k!}\Psi_I^*$ cancels. Otherwise, there is some $\sigma \in S_k$ and some increasing multi-index J such that:

$$\frac{a_I}{k!}\Psi_I = \frac{a_I}{k!}\Psi_{J^\sigma} = \frac{a_I \operatorname{sgn}(\sigma)}{k!}\Psi_J.$$

By collecting terms, we get that $T = \sum_{J \text{ increasing}} c_J \Phi_J$ where each $c_J \in F$.

This shows that the Ψ_J span all of $\mathcal{A}^k(V)$. Next we show that they form a basis. Suppose $T = \sum_{J \text{ increasing}} c_J \Phi_J = 0$.

Then by part (c) of the last proposition, we know that if $I = (i_1, \ldots, i_k)$ is an increasing multi-index, then:

$$0 = T(u_{i_1}, \ldots, u_{i_k}) = C_I$$

So, all the C_I are equal to 0. ■

Corollary: If F has characteristic 0 or greater than k, then $\mathcal{A}^k(V)$ has dimension $\binom{n}{k}$.

Corollary 2: If F has characteristic 0 or greater than $k \geq n$, then any alternating n-tensor on V is a scalar multiple of a determinant function. Also, there are no nontrivial alternating m-tensors where $n < m \leq k$.

Exercise 1.4.ix: Suppose $A : V \to W$ is a linear map. Then if $T \in \mathcal{A}^k(W)$, we have that $A^\dagger T \in \mathcal{A}^k(V)$. Hence, the pullback operation maps alternating tensors to alternating tensors.

Proof:
Suppose $\sigma \in S_k$. Then for any $v_1, \ldots, v_k \in V$, we have that:

$$\begin{aligned}(A^\dagger T)^\sigma(v_1, \ldots, v_k) &= A^\dagger T(v_{\sigma^{-1}(1)}, \ldots, v_{\sigma^{-1}(k)}) \\ &= T(Av_{\sigma^{-1}(1)}, \ldots, Av_{\sigma^{-1}(k)}) \\ &= T^\sigma(Av_1, \ldots, Av_k) \\ &= \operatorname{sgn}(\sigma)T(Av_1, \ldots, Av_k) = \operatorname{sgn}(\sigma)A^\dagger T(v_1, \ldots, v_k)\end{aligned}$$

Hence, $(A^\dagger T)^\sigma = \operatorname{sgn}(\sigma)A^\dagger T$. This proves that $A^\dagger T$ is alternating. ■

Exercise 1.4.x: Additionally to the last exercise, we have that if $T \in \mathcal{L}^k(V)$, then $A^\dagger(\operatorname{Alt}(T)) = \operatorname{Alt}(A^\dagger T)$.

Proof:
If $v_1, \ldots, v_k \in V$, then:

$$\begin{aligned}\operatorname{Alt}(A^\dagger T)(v_1, \ldots, v_k) &= \sum_{\tau \in S_k} \operatorname{sgn}(\tau)(A^\dagger T)^\tau(v_1, \ldots, v_k) \\ &= \sum_{\tau \in S_k} \operatorname{sgn}(\tau) A^\dagger T(v_{\tau^{-1}(1)}, \ldots, v_{\tau^{-1}(k)}) \\ &= \sum_{\tau \in S_k} \operatorname{sgn}(\tau) T(Av_{\tau^{-1}(1)}, \ldots, Av_{\tau^{-1}(k)})\end{aligned}$$

$$= \sum_{\tau \in S_k} \operatorname{sgn}(\tau) T^{\tau}(Av_1, \ldots, Av_k)$$
$$= \operatorname{Alt}(T)(Av_1, \ldots, Av_k) = A^{\dagger}(\operatorname{Alt}(T))(v_1, \ldots, v_k)$$

This shows that $\operatorname{Alt}(A^{\dagger}T) = A^{\dagger}(\operatorname{Alt}(T))$. ■

The space $\Lambda^k(V^*)$:

If $k > 1$, a decomposable k-tensor $\ell_1 \otimes \cdots \otimes \ell_k$ with each $\ell_i \in V^*$ is called redundant if $\ell_i = \ell_{i+1}$ for some index i. We let $\mathcal{I}^k(V)$ be the span of all redundant k-tensors.

If $k = 1$, we define $\mathcal{I}^1(V) := \{0\} \subseteq \mathcal{L}^1(V)$.

Also if $k = 0$, we define $\mathcal{I}^0(V) := \{0\} \subseteq F$.

Proposition 1.5.2: Suppose F has characteristic $\neq 2$. If $T \in \mathcal{I}^k(V)$, then $\operatorname{Alt}(T) = 0$. In other words, $\mathcal{I}^k(V) \subseteq \ker(\operatorname{Alt})$.

Proof:

If $T \in \mathcal{I}^k(V)$, then we know there are redundant decomposable k-tensors $T_1, \ldots, T_m$ as well as scalars $c_1, \ldots, c_m \in F$ such that $T = \sum_{j=1}^m c_j T_j$. Then since $\operatorname{Alt}(T) = \sum_{j=1}^m c_j \operatorname{Alt}(T_j)$, all we need to do now is show that $\operatorname{Alt}(T_j) = 0$ for every j.

Since T_j is a redundant decomposable k-tensor, we know that $T_j = \ell_1 \otimes \cdots \otimes \ell_k$ where $\ell_i = \ell_{i+1}$ for some $1 \leq i < k$. In turn, if $\tau_{i,i+1} \in S_k$ is the transposition of i and $i+1$, we have that $(T_j)^{\tau_{i,i+1}} = T_j$ and $\operatorname{sgn}(\tau_{i,i+1}) = -1$. Hence:

$$\operatorname{Alt}(T_j) = \operatorname{Alt}((T_j)^{\tau_{i,i+1}}) = \operatorname{sgn}(\tau_{i,i+1})\operatorname{Alt}(T_j) = -\operatorname{Alt}(T_j)$$

This implies that $\operatorname{Alt}(T_j) = 0$. ■

Proposition 1.5.3: If $T \in \mathcal{I}^r(V)$ and $T' \in \mathcal{L}^s(V)$, then $T \otimes T'$ and $T' \otimes T$ are in $\mathcal{I}^{r+s}(V)$.

Proof:

The argument for $T' \otimes T$ being in $\mathcal{I}^{r+s}(V)$ is mostly identical to the argument for $T \otimes T'$ being in $\mathcal{I}^{r+s}(V)$. So, I'll focus only on proving the latter.

To start off, like before we know that there are redundant decomposable r-tensors $T_1, \ldots, T_m$ as well as scalars $c_1, \ldots, c_m \in F$ such that $T = \sum_{j=1}^m c_j T_j$. Hence, it suffices to show that $T_j \otimes T' \in \mathcal{I}^{r+s}(V)$ for all $1 \leq j \leq m$ since:

$$T \otimes T' = (\textstyle\sum_{j=1}^m c_j T_j) \otimes T' = \sum_{j=1}^m c_j (T_j \otimes T')$$

Fortunately, by writing $T' = \sum_I d_I u_I^*$, we can see that:

$$T_j \otimes T' = T_j \otimes (\sum_I d_I u_I^*) = \sum_I d_I (T_j \otimes u_I^*)$$

Now since both u_I^* and T_j are decomposable and T_j is redundant, we can easily see that $T_j \otimes u_I^*$ is decomposable and redundant. It follows that $T_j \otimes T' \in \mathcal{I}^{r+s}(V)$. ■

Proposition 1.5.4: Suppose F has characteristic $\neq 2$. If $T \in \mathcal{L}^k(V)$ and $\sigma \in S_k$, then $T^{\sigma} = \operatorname{sgn}(\sigma) T + S$ where $S \in \mathcal{I}^k(V)$.

Proof:
Hopefully you're getting use to this trick. It suffices to assume T is decomposable. After all, after writing $T = \sum_I c_I u_I^*$, if we can show for all multi-indexes I that $(u_I^*)^\sigma = \mathrm{sgn}(\sigma)u_I^* + S_I$ where $S_I \in \mathcal{I}^k(V)$, then we can set $S = \sum_I c_I S_I \in \mathcal{I}^k(V)$ and have that:

$$T^\sigma = \sum_I c_I(u_I^*)^\sigma = \mathrm{sgn}(\sigma)\sum_I c_I u_I^* + \sum_I S_I = \mathrm{sgn}(\sigma)T + S$$

So suppose $T = \ell_1 \otimes \cdots \otimes \ell_k$. Then given $\sigma \in S^k$, we can write $\sigma = \tau_1 \ldots \tau_m$ as the product of m many transpositions of adjacent pairs of numbers in $\{1, \ldots, k\}$.

(by adjacent I mean a pair $\{j, j+1\}$ where $1 \leq j < k$...)

We shall induct on m. First assume $m = 1$. Thus $\sigma = \tau_{j,j+1}$ for some $1 \leq j < k$ and hence $\mathrm{sgn}(\sigma) = -1$. Also:

$$\begin{aligned} T^\sigma - \mathrm{sgn}(\sigma)T &= T^\sigma + T \\ &= (\ell_1 \otimes \cdots \otimes \ell_{j-1} \otimes \ell_{j+1} \otimes \ell_j \otimes \ell_{j+2} \otimes \cdots \otimes \ell_k) + (\ell_1 \otimes \cdots \otimes \ell_k) \\ &= (\ell_1 \otimes \cdots \otimes \ell_{j-1}) \otimes ((\ell_{j+1} \otimes \ell_j) + (\ell_j \otimes \ell_{j+1})) \otimes (\ell_{j+2} \otimes \cdots \otimes \ell_k) \end{aligned}$$

Now note that:

$$(\ell_j + \ell_{j+1}) \otimes (\ell_j + \ell_{j+1}) = (\ell_j \otimes \ell_j) + (\ell_j \otimes \ell_{j+1}) + (\ell_{j+1} \otimes \ell_j) + (\ell_{j+1} \otimes \ell_{j+1}).$$

Therefore:

$$\begin{aligned} T^\sigma - \mathrm{sgn}(\sigma)T = (\ell_1 \otimes \cdots \otimes \ell_{j-1}) &\otimes (\ell_j + \ell_{j+1}) \otimes (\ell_j + \ell_{j+1}) \otimes (\ell_{j+1} \otimes \cdots \otimes \ell_k) \\ &- (\ell_1 \otimes \cdots \otimes \ell_{j-1}) \otimes \ell_j \otimes \ell_j \otimes (\ell_{j+2} \otimes \cdots \otimes \ell_k) \\ &- (\ell_1 \otimes \cdots \otimes \ell_{j-1}) \otimes \ell_{j+1} \otimes \ell_{j+1} \otimes (\ell_{j+2} \otimes \cdots \otimes \ell_k) \end{aligned}$$

Hence $T^\sigma - \mathrm{sgn}(\sigma)T \in \mathcal{I}^k(V)$ and we are done with this case.

Now suppose $m > 1$. Then $\sigma = \tau_{j,j+1}\sigma'$ where σ' is the product of $m-1$ transpositions. By induction, we know that there exists $S_1 \in \mathcal{I}^k(V)$ such that:

$$T^\sigma = (T^{\tau_{j,j+1}})^{\sigma'} = \mathrm{sgn}(\sigma')T^{\tau_{j,j+1}} + S_1$$

Also by our base case, there is $S_2 \in \mathcal{I}^k(V)$ such that $T^{\tau_{j,j+1}} = \mathrm{sgn}(\tau_{j,j+1})T + S_2$. Then setting $S = \mathrm{sgn}(\tau_{j,j+1})S_2 + S_1$, we have that $S \in \mathcal{I}^K(V)$ and:

$$T^\sigma = \mathrm{sgn}(\sigma')(\mathrm{sgn}(\tau_{j,j+1})T + S_2) + S_1 = \mathrm{sgn}(\sigma)T + S. \blacksquare$$

<u>Corollary 1.5.6:</u> Suppose F has characteristic $\neq 2$. If $T \in \mathcal{L}^k(V)$, then $\mathrm{Alt}(T) = k!T + S$ where $S \in \mathcal{I}^k(V)$.

Proof:
Given any $\tau \in S_k$, let $S_\tau \in \mathcal{I}^k(V)$ be such that $T^\tau = \mathrm{sgn}(\tau)T + S_\tau$. Then $S := \sum_{\tau \in S_k} \mathrm{sgn}(\tau)S_\tau \in \mathcal{I}^k(V)$ and:

$$\mathrm{Alt}(T) = \sum_{\tau \in S_k} \mathrm{sgn}(\tau)T^\tau = \sum_{\tau \in S_k} (\mathrm{sgn}(\tau))^2 T + \sum_{\tau \in S_k} \mathrm{sgn}(\tau)S_\tau = k!T + S. \blacksquare$$

Corollary 1.5.8: Let $k \geq 1$. Then let V be a vector space over a field F of characteristic 0 or $> \max(k, 2)$. Then:

$$\mathcal{I}^k(V) = \ker(\mathrm{Alt} : \mathcal{L}^k(V) \to \mathcal{A}^k(V))$$

Proof:
We already know from proposition 1.5.2 that $\mathcal{I}^k(V) \subseteq \ker(\mathrm{Alt})$. To prove the reverse relation, suppose $T \in \mathcal{L}^k(V)$ satisfies that $\mathrm{Alt}(T) = 0$. Then based on the previous corollary, we know there exists $S \in \mathcal{I}^k(V)$ such that $-\frac{1}{k!}S = T - \mathrm{Alt}(T)$. Hence $T \in \mathcal{I}^k(V)$. ■

Theorem 1.5.9: Suppose F is a field of characteristic 0 or $> \max(k, 2)$. Then any element $T \in \mathcal{L}^k(V)$ can be written uniquely as a sum $T_1 + T_2$ where $T_1 \in \mathcal{A}^k(V)$ and $T_2 \in \mathcal{I}^k(V)$. I.e, $\mathcal{L}^k(V) = \mathcal{A}^k(V) \oplus \mathcal{I}^k(V)$.

Proof:
Let $W \in \mathcal{I}^k(V)$ satisfy that $\mathrm{Alt}(T) = k!T + W$. Then set $T_1 = \frac{1}{k!}\mathrm{Alt}(T)$ and $T_2 = -\frac{1}{k!}W$. Then clearly $T = T_1 + T_2$ with $T_1 \in \mathcal{A}^k(V)$ and $T_2 \in \mathcal{I}^k(V)$.

Next, to prove uniqueness suppose $T_1' + T' = T$ with $T_1' \in \mathcal{A}^k(V)$ and $T_2' \in \mathcal{I}^k(V)$. Then $T_1 - T_1' \in \mathcal{A}^k(V)$, $T_2 - T_2' \in \mathcal{I}^k(V)$, and $(T_1 - T_1') + (T_2 - T_2') = 0$. So:

$$0 = \mathrm{Alt}(0) = \mathrm{Alt}((T_1 - T_1') + (T_2 - T_2')) = k!(T_1 - T_1')$$

Hence $T_1 = T_1'$ and it easily follows $T_2 = T_2'$. ■

Let $k \geq 0$. Let V be a finite dimensional vector space over a field F of characteristic 0 or $> \max(k, 2)$. Then we define:

$$\Lambda^k(V^*) \coloneqq \mathcal{L}^k(V)/\mathcal{I}^k(V)$$

By the first isomorphism theorem along with the previous theorem, we have that $\Lambda^k(V^*) \cong \mathcal{A}^k(V)$.

8/8/2025

Here is a tangent about symmetric tensors. For this section, suppose V is an n-dimensional vector space over a field F with characteristic $\neq 2$.

A tensor $T \in \mathcal{L}^k(V)$ is symmetric if $T^\sigma = T$ for all $\sigma \in S_k$. We denote the space of symmetric tensors $\mathcal{S}^k(V)$.

You can show by the same reasoning as with $\mathcal{A}^k(V)$ that $\mathcal{S}^k(V)$ is a vector subspace.

Exercise 1.5.iii: Suppose F has characteristic 0 or $> k$. Then if T is a symmetric k-tensor and $k \geq 2$, we have that $T \in \mathcal{I}^k(V)$.

Proof:
Let $\sigma \in S_k$ be an odd permutation. Then by proposition 1.4.17:

$$\mathrm{Alt}(T) = \mathrm{Alt}(T^\sigma) = \mathrm{sgn}(\sigma)\mathrm{Alt}(T) = -\mathrm{Alt}(T).$$

The only way this is possible is if $\mathrm{Alt}(T) = 0$. Hence $T \in \ker(\mathrm{Alt})$, and by theorem 1.5.8 that means that $T \in \mathcal{I}^k(V)$. ■

We define a symmetrization operator as follows. Given $T \in \mathcal{L}^k(V)$, define:

$$\mathrm{Sym}(T) := \sum_{\sigma \in S_k} T^\sigma$$

Then like in proposition 1.4.17, we can show that given any $T \in \mathcal{L}^k(V)$ and $\sigma \in S_k$:

(a) $(\mathrm{Sym}(T))^\sigma = \mathrm{Sym}(T)$ (i.e. $\mathrm{Sym}(T) \in \mathcal{S}^k(V)$...)

(b) If $T \in \mathcal{S}^k(V)$, then $\mathrm{Sym}(T) = k!T$

(c) $\mathrm{Sym}(T^\sigma) = \mathrm{Sym}(T)$

(d) $\mathrm{Sym} : \mathcal{L}^k(V) \to \mathcal{A}^k(V)$ is an linear map (which is surjective so long as $k! \neq 0$ in the field F...).

Supposing $k! \neq 0$ in F, then by following a process very similar to what we did with the alternation operation, we can construct a basis for the symmetric tensors:

$$\{\Phi_I^* : I \text{ is a non-decreasing multi-index}\}.$$

Note: By non-decreasing the textbook means that $I = (i_1, \ldots, i_k)$ satisfies that $i_1 \leq i_2 \leq \cdots \leq i_k$.

I'm bored and won't do that construction here. But the important point is that this means $\mathcal{S}^k(V)$ has the same number of dimensions as there are non-decreasing multi-indexes of n of length k. And since there are $\binom{n+k-1}{k}$ ways of picking k elements of the set $\{1, \ldots, n\}$ when you allow yourself to pick the same element multiple times, this means that $\dim(\mathcal{S}^k(V)) = \binom{n+k-1}{k}$.

Side note: If $k! \neq 0$ in F, then we have already shown that $\dim(\mathcal{I}^k(V)) = n^k - \binom{n}{k}$. Since $\mathcal{S}^k(V) \subseteq \mathcal{I}^k(V)$, this shows that $\mathcal{S}^2(V) = \mathcal{I}^2(V)$. That said, we don't in general have that $\dim(\mathcal{I}^k(V)) = \dim(\mathcal{S}^k(V))$ when $k > 2$.

Next, here's some other miscellaneous results.

Exercise 1.5.vii: Suppose F has characteristic 0 or $> \max(k, 2)$. Then if $T \in \mathcal{I}^k(V)$, we have that $T^\sigma \in \mathcal{I}^k(V)$ for all $\sigma \in S_k$.

Proof:
Since $T \in \mathcal{I}^k(V)$, we know that: $\mathrm{Alt}(T^\sigma) = \mathrm{sgn}(\sigma)\mathrm{Alt}(T) = 0$. Therefore, $T^\sigma \in \ker(\mathrm{Alt})$, and by corollary 1.5.8 we know that $\ker(\mathrm{Alt}) = \mathcal{I}^k(V)$. ■

Corollary / Exercise 1.5.v: Let $k \geq 2$ and suppose F has characteristic 0 or $> k$. Then if $T \in \mathcal{L}^{k-2}(V)$ and $\ell \in V^*$, we have that $\ell \otimes T \otimes \ell \in \mathcal{I}^k(V)$.

Proof:
There is a permutation $\sigma \in S_k$ satisfying that $(\ell \otimes T \otimes \ell)^\sigma = (\ell \otimes \ell) \otimes T$. Then by proposition 1.5.3, we have that $(\ell \otimes \ell) \otimes T \in \mathcal{I}^k(V)$. And finally, by applying the last exercise we have that $\ell \otimes T \otimes \ell = ((\ell \otimes \ell) \otimes T)^{\sigma^{-1}} \in \mathcal{I}^k(V)$. ■

Corollary / Exercise 1.5.vi: Let $k \geq 2$ and suppose F has characteristic 0 or $> k$. Then if $T \in \mathcal{L}^{k-2}(V)$ and $\ell_1, \ell_2 \in V^*$, we have that $(\ell_1 \otimes T \otimes \ell_2) + (\ell_2 \otimes T \otimes \ell_1) \in \mathcal{I}^k(V)$.

Proof:
Apply the last exercise plus the fact that:

$$\begin{aligned}(\ell_1 \otimes T \otimes \ell_2) + (\ell_2 \otimes T \otimes \ell_1) = ((\ell_1 + \ell_2) \otimes T \otimes (\ell_1 + \ell_2)) \\ - (\ell_1 \otimes T \otimes \ell_1) - (\ell_2 \otimes T \otimes \ell_2). \blacksquare\end{aligned}$$

8/9/2025

The Wedge Product:
In this section, we'll suppose V is an n-dimensional vector space over a field F of characteristic 0. Also, we shall for each k define the map $\pi : \mathcal{L}^k(V) \to \Lambda^k(V^*)$ such that $\pi(T) = T + \mathcal{I}^k(V)$ for all $T \in \mathcal{L}^k(V)$.

Suppose for each $i \in \{1, 2\}$ we have $\omega_i \in \Lambda^{k_i}(V^*)$. Then if for each i we are given $T_i \in \mathcal{L}^{k_i}(V)$ satisfying that $\pi(T_i) = \omega_i$, we define:

$$\omega_1 \wedge \omega_2 \coloneqq \pi(T_1 \otimes T_2).$$

Claim 1.6.3: The wedge product is well defined.

Proof:
Suppose for each $i \in \{1, 2\}$ that we also have $T_i' \in \mathcal{L}^{k_i}(V)$ satisfying that $\pi(T_i') = \pi(T_i) = \omega_i$. Then for each i there exists $W_i \in \mathcal{I}^k(V)$ such that $T_i' = T_i + W_i$. Hence:

$$\pi(T_1' \otimes T_2') = \pi((T_1 \otimes T_2) + (T_1 \otimes W_2) + (W_1 \otimes T_2) + (W_1 \otimes W_2))$$

Then by applying proposition 1.5.3, we know that:

$$(T_1 \otimes W_2) + (W_1 \otimes T_2) + (W_1 \otimes W_2) \in \mathcal{I}^k(V)$$

Hence, $\pi(T_1' \otimes T_2') = \pi(T_1 \otimes T_2)$. $\blacksquare$

More generally, if for each $i \in \{1, \ldots, m\}$ we have $\omega_i \in \Lambda^{k_i}(V^*)$ and $T_i \in \mathcal{L}^{k_i}(V)$ satsifying that $\pi(T_i) = \omega_i$, then we define:

$$\omega_1 \wedge \cdots \wedge \omega_m \coloneqq \pi(T_1 \otimes \cdots \otimes T_m)$$

This is well defined for basically the same reasoning as before, although to avoid some overly long expressions, it suffices to replace only one tensor at a time.

Claim: Given any $m \geq 3$, we have that:

$$\omega_1 \wedge (\omega_2 \wedge \cdots \wedge \omega_m) = \omega_1 \wedge \cdots \wedge \omega_m = (\omega_1 \wedge \cdots \wedge \omega_{m-1}) \wedge \omega_m$$

Proof:
If for each $i \in \{1, \ldots, m\}$ we have some $T_i \in \mathcal{L}^{k_i}(V)$ satisfying that $\pi(T_i) = \omega_i$, then $\pi(T_2 \otimes \cdots \otimes T_m) = \omega_2 \wedge \cdots \wedge \omega_m$ and $\pi(T_1 \otimes \cdots \otimes T_{m-1}) = \omega_1 \wedge \cdots \wedge \omega_{m-1}$.
In turn:

$$\begin{aligned}\omega_1 \wedge \cdots \wedge \omega_m &= \pi(T_1 \otimes \cdots \otimes T_m) \\ &= \pi(T_1 \otimes (T_2 \otimes \cdots \otimes T_m)) = \omega_1 \wedge (\omega_2 \wedge \cdots \wedge \omega_m) \\ &= \pi((T_1 \otimes \cdots \otimes T_{m-1}) \otimes T_m) = (\omega_1 \wedge \cdots \wedge \omega_{m-1}) \wedge \omega_m\end{aligned}$$

Corollary: The wedge product is associative and we get the same result no matter how we use parentheses to group together the ω_i.

Proof:

Suppose we have $\omega_1, \ldots, \omega_m$. If $m = 3$, then we're already done by the last claim. Meanwhile, for $m > 3$ is suffices due to the strong inductive hypothesis on m to show that for any $r \in \{1, \ldots, m-1\}$:

$$\omega_1 \wedge \cdots \wedge \omega_m = (\omega_1 \wedge \cdots \wedge \omega_r) \wedge (\omega_{r+1} \wedge \cdots \wedge \omega_m)$$

Luckily, note that by the previous claim as well as our inductive hypothesis:

$$\begin{aligned}(\omega_1 \wedge \cdots \wedge \omega_r) \wedge (\omega_{r+1} \wedge \cdots \wedge \omega_m) &= ((\omega_1 \wedge \cdots \wedge \omega_r) \wedge \omega_{r+1} \wedge \cdots \wedge \omega_{m-1}) \wedge \omega_m \\ &= (\omega_1 \wedge \cdots \wedge \omega_{m-1}) \wedge \omega_m \\ &= \omega_1 \wedge \cdots \wedge \omega_m. \blacksquare\end{aligned}$$

Here are some other properties of the wedge product which I'm too bored to properly prove:

- If $\lambda \in F$, then $\lambda(\omega_1 \wedge \omega_2) = (\lambda\omega_1) \wedge \omega_2 = \omega_1 \wedge (\lambda\omega_2)$.
- $(\omega_1 + \omega_2) \wedge \omega_3 = (\omega_1 \wedge \omega_3) + (\omega_2 \wedge \omega_3)$
- $\omega_1 \wedge (\omega_2 + \omega_3) = (\omega_1 \wedge \omega_2) + (\omega_1 \wedge \omega_3)$

Side note: if we were instead writing the definition of the wedge product in terms of alternating tensors, we'd be defining:

$$T_1 \wedge \cdots \wedge T_m = \frac{1}{(k_1 + \ldots + k_m)!}\mathrm{Alt}(T_1 \otimes \cdots \otimes T_m).$$

Hopefully its obvious why this definition is inferior.

Note that since $\mathcal{I}^1(V) = \{0\}$, we can just identify $\Lambda^1(V^*) = V^*$. Then given $\ell_1, \ldots, \ell_k \in V^* = \Lambda^1(V^*)$, we say that $\omega = \pi(\ell_1 \otimes \cdots \otimes \ell_k) = \ell_1 \wedge \cdots \wedge \ell_k$ is a decomposable element of $\Lambda^k(V^*)$.

Claim: For any $\sigma \in S_k$, we have: $\ell_{\sigma(1)} \wedge \cdots \wedge \ell_{\sigma(k)} = \mathrm{sgn}(\sigma)\ell_1 \wedge \cdots \wedge \ell_k$.

Proof:

Lemma: If $T \in \mathcal{L}^k(V)$ and $\sigma \in S_k$, then $\pi(T^\sigma) = \mathrm{sgn}(\sigma)\pi(T)$.

This is just a consequence of proposition 1.5.4.

As a result of that lemma:

$$\begin{aligned}\ell_{\sigma(1)} \wedge \cdots \wedge \ell_{\sigma(k)} &= \pi(\ell_{\sigma(1)} \otimes \cdots \otimes \ell_{\sigma(k)}) \\ &= \pi((\ell_1 \otimes \cdots \otimes \ell_k)^\sigma) \\ &= \mathrm{sgn}(\sigma)\pi(\ell_1 \otimes \cdots \otimes \ell_k) \\ &= \mathrm{sgn}(\sigma)\ell_1 \wedge \cdots \wedge \ell_k\end{aligned}$$

As a corollary, given any $\ell_1, \ell_2 \in V^*$ we have that $\ell_1 \wedge \ell_2 = -\ell_2 \wedge \ell_1$. Also, given $\ell_1, \ell_2, \ell_3 \in V^*$, we have:

$$\begin{aligned}\ell_1 \wedge \ell_2 \wedge \ell_3 &= -\ell_2 \wedge \ell_1 \wedge \ell_3 = \ell_2 \wedge \ell_3 \wedge \ell_1 \\ &= -\ell_1 \wedge \ell_3 \wedge \ell_2 = \ell_3 \wedge \ell_1 \wedge \ell_2\end{aligned}$$

Let $u_1, \dots, u_n$ be a basis for V and let $u_1^*, \dots, u_n^*$ be the corresponding dual basis. Then the collection of $u_{i_1}^* \wedge \cdots \wedge u_{i_k}^*$ such that $I = (i_1, \dots, i_k)$ is an increasing multi-index forms a basis for $\Lambda^k(V^*)$.

Proof:
Recall that when defining $u_I^* = u_{i_1}^* \otimes \cdots \otimes u_{i_k}$ for a multi-index $I = (i_1, \dots, i_k)$, we then have that the $\Psi_I := \mathrm{Alt}(u_I^*)$ where I is increasing form a basis of $\mathcal{A}^k(V)$. It follows that each $\pi(\Psi_I)$ where I is increasing is a basis vector of $\Lambda^k(V^*)$. But note that:

$$\pi(\Psi_I) = \pi\left(\sum_{\tau \in S_k} \mathrm{sgn}(\tau)(u_I^*)\tau\right) = \sum_{\tau \in S_k} \mathrm{sgn}(\tau)\pi((u_I^*)^\tau) = \sum_{\tau \in S_k} (\mathrm{sgn}(\tau))^2 \pi(u_I^*) = k!\pi(u_I^*)$$

So, the $\pi(u_I^*) = u_{i_1} \wedge \cdots \wedge u_{i_k}$ also form a basis for $\Lambda^k(V^*)$.

This now let's us prove the following general result:

Theorem 1.6.10: If $\omega_1 \in \Lambda^r(V^*)$ and $\omega_2 \in \Lambda^s(V^*)$, then $\omega_1 \wedge \omega_2 = (-1)^{rs}\omega_2 \wedge \omega_1$.

Proof:
Express $\omega_1 = \sum_I c_I u_{i_1}^* \wedge \cdots \wedge u_{i_r}^*$ and $\omega_2 = \sum_J d_J u_{j_1}^* \wedge \cdots \wedge u_{j_s}^*$.

Then we have that:

$$\begin{aligned}
\omega_1 \wedge \omega_2 &= \sum_{I,J} c_I d_J (u_{i_1}^* \wedge \cdots \wedge u_{i_r}^* \wedge u_{j_1}^* \wedge \cdots \wedge u_{j_s}^*) \\
&= \sum_{I,J} c_I d_J (-1)^{rs} (u_{j_1}^* \wedge \cdots \wedge u_{j_s}^* \wedge u_{i_1}^* \wedge \cdots \wedge u_{i_r}^*) \\
&= (-1)^{rs} \sum_{I,J} d_J c_I (u_{j_1}^* \wedge \cdots \wedge u_{j_s}^* \wedge u_{i_1}^* \wedge \cdots \wedge u_{i_r}^*) = (-1)^{rs}\omega_2 \wedge \omega_1.
\end{aligned}$$

One more note I'd like to make is that we can identify $\Lambda^0(V^*)$ and F. Then if $\omega \in \Lambda^k(V^*)$ and $\lambda \in F$, we have that $\lambda \wedge \omega = \lambda\omega = \omega \wedge \lambda$.

8/10/2025

Before moving onto the next section of the book, I'm going to do a few of the exercises.

Exercise 1.6.iii: Given $\omega \in \Lambda^r(V^*)$, we define $\omega^1 := \omega$ and $\omega^k := \omega \wedge \omega^{k-1} \in \Lambda^{rk}(V^*)$ for all $k > 1$. In other words, ω^k is the k-fold wedge product of ω with itself.

(A) If r is odd, then $\omega^k = 0$ for all $k > 1$.

Proof:
By an easy application of theorem 1.6.10, we have that:

$$\omega^k = \omega \wedge \omega^{k-1} = (-1)^{r \cdot r^{k-1}} \omega^{k-1} \wedge \omega = (-1)^{r^k}\omega^k$$

But r^k is odd if r is odd. Then in turn, $\omega^k = -\omega^k$. The only way this is possible is if $\omega^k = 0$.

(B) If ω is decomposable, then $\omega^k = 0$ for all $k > 1$.

Proof:
For the ease of notation we'll $\omega^0 = 1 \in F$. Now if $\omega = \ell_1 \wedge \cdots \wedge \ell_r$, then by just swapping two occurences of ℓ_1, we have that:

$$\begin{aligned}\omega^k &= \ell_1 \wedge \cdots \wedge \ell_r \wedge \ell_1 \wedge \cdots \wedge \ell_r \wedge \omega^{k-2} \\ &= (-1)\ell_1 \wedge \cdots \wedge \ell_r \wedge \ell_1 \wedge \cdots \wedge \ell_r \wedge \omega^{k-2} = -\omega^k\end{aligned}$$

This implies $\omega^k = 0$.

Exercise 1.6.iv: If $\omega, \mu \in \Lambda^r(V^*)$, then:

$$(\omega + \mu)^k = \sum_{i=0}^{k} \binom{k}{i} \omega^i \wedge \mu^{k-i}.$$

This is obvious so I'm skipping this problem. I just wanted to write out the result.

The interior Product:
All the assumptions about V and F made yesterday still apply and you should keep assuming them until I tell you to stop (cause I don't want to keep writing this shtick).

Given $T \in \mathcal{L}^k(V)$ where $k > 1$ and $v \in V$, we define the $(k-1)$-tensor:

$$\iota_v T(v_1, \ldots, v_{k-1}) \coloneqq \sum_{r=1}^{k} (-1)^{r-1} T(v_1, \ldots, v_{r-1}, v, v_r, \ldots, v_{k-1})$$

Also if $\lambda \in \mathcal{L}^0(V) = F$, we define $\iota_v \lambda = 0$ for all $v \in V$.

Note that if $v = c_1 v_1 + c_2 v_2$ and $T = d_1 T_1 + d_2 T_2$, then:

$$\iota_v T = c_1 \iota_{v_1} T + c_2 \iota_{v_2} T \text{ and } \iota_v T = d_1 \iota_v T_1 + d_2 \iota_v T_2.$$

Also, if $T = \ell_1 \otimes \cdots \otimes \ell_k$ where each $\ell_i \in V^*$, then when writing $\hat{\ell_r}$ to mean that we are deleting ℓ_r from that term of the expression, we have that:

$$\iota_v T = \sum_{r=1}^{k} (-1)^{r-1} \ell_r(v) \ell_1 \otimes \cdots \otimes \hat{\ell_r} \otimes \cdots \otimes \ell_k$$

Slightly less obviously, if $T_1 \in \mathcal{L}^p(V)$ and $T_2 \in \mathcal{L}^q(V)$, we have that:

$$\iota_v(T_1 \otimes T_2) = (\iota_v T_1) \otimes T_2 + (-1)^p T_1 \otimes (\iota_v T_2).$$

Lemma 1.7.8: Let V be a vector space and $T \in \mathcal{L}^k(V)$ where $k \geq 1$. Then for all $v \in V$, $\iota_v(\iota_v(T)) = 0$.

Proof:
By linearity it suffices to prove this statement for decomposable T. Also, this statement is trivial when $k = 1$. So, we can proceed by induction, assuming that the theorem holds for $T \in \mathcal{L}^r(V)$ where $r < k$. Then after expressing $T = T' \otimes \ell$ where $T' \in \mathcal{L}^{k-1}(V)$ and $\ell \in V^*$, we have that:

$$\iota_v T = \iota_v(T' \otimes \ell) = (\iota_v T') \otimes \ell + (-1)^{k-1} T' \otimes (\iota_v T')$$
$$= (\iota_v T') \otimes \ell + (-1)^{k-1} \ell(v) T'$$

By induction $\iota_v(\iota_v T') = 0$. Combining that with the above reasoning shows:

$$\begin{aligned}\iota_v(\iota_v T) &= \iota_v((\iota_v T') \otimes \ell + (-1)^{k-1}\ell(v)T') \\ &= \iota_v((\iota_v T') \otimes \ell) + (-1)^{k-1}\ell(v)\iota_v(T') \\ &= \big(\iota_v(\iota_v T') \otimes \ell + (-1)^{k-2}(\iota_v T') \otimes (\iota_v \ell)\big) + (-1)^{k-1}\ell(v)\iota_v(T') \\ &= 0 + (-1)^{k-2}\ell(v)(\iota_v T') + (-1)^{k-1}\ell(v)\iota_v(T') = 0. \blacksquare\end{aligned}$$

Corollary: If $v_1, v_2 \in V$ and $T \in \mathcal{L}^k(V)$, then $\iota_{v_1}(\iota_{v_2} T) = -\iota_{v_2}(\iota_{v_1} T)$.

Proof:

We know from the prior lemma that:

$$\iota_{v_1+v_2}(\iota_{v_1+v_2} T) = 0$$

Therefore:

$$0 + \iota_{v_1}(\iota_{v_2} T) = \iota_{v_1}(\iota_{v_1+v_2} T) = -\iota_{v_2}(\iota_{v_1+v_2} T) = -\iota_{v_2}(\iota_{v_1} T) - 0$$

Lemma 1.7.11: If $T \in \mathcal{L}^k(V)$ is redundant, then so is $\iota_v T$.

Proof:

Write $T = T_1 \otimes \ell \otimes \ell \otimes T_2$ where $\ell \in V^*$, $T_1 \in \mathcal{L}^p(V)$, and $T_2 \in \mathcal{L}^q(V)$. Then:

$$\iota_v T = \iota_v(T_1) \otimes \ell \otimes \ell \otimes T_2 + (-1)^p T_1 \otimes \iota_v(\ell \otimes \ell) \otimes T_2 + (-1)^{p+2} T_1 \otimes \ell \otimes \ell \otimes \iota_v(T_2)$$

Now the first and third terms are obvious redundant. Meanwhile, the second term cancels because $\iota_v(\ell \otimes \ell) = \ell(v)\ell - \ell(v)\ell = 0$. $\blacksquare$

Corollary: If $T \in \mathcal{I}^k(V)$, then $\iota_v T \in \mathcal{I}^{k-1}(V)$.

Now we define the interior product operator ι_v on $\Lambda^k(V^*)$. If π is the projection of $\mathcal{L}^k(V)$ onto $\Lambda^k(V^*)$ and $\omega = \pi(T) \in \Lambda^k(V^*)$, then we define:

$$\iota_v \omega \coloneqq \pi(\iota_v T) \in \Lambda^{k-1}(V^*).$$

This is well defined since by the previous corollary, if both T and T' satisfy that $\pi(T) = \pi(T') = \omega$, then there is some tensor $S \in \mathcal{I}^{k-1}(V)$ such that $\iota_v T = \iota_v T' + S$.

It is easily shown then that if $v_1, v_2, v \in V$, $\omega, \omega_1 \in \Lambda^p(V^*)$, and $\omega_2 \in \Lambda^q(V^*)$, then:

- $\iota_{v_1+v_2}\omega = \iota_{v_1}\omega + \iota_{v_2}\omega$;
- $\iota_v(\lambda_1\omega_1 + \lambda_2\omega_2) = \lambda_1\iota_v\omega_1 + \lambda_2\iota_v\omega_2$ (where $\lambda_1, \lambda_2 \in F$);
- $\iota_v(\omega_1 \wedge \omega_2) = (\iota_v\omega_1) \wedge \omega_2 + (-1)^p \omega_1 \wedge (\iota_v\omega_2)$.

Also, if you squint you can see that $\iota_v(\iota_v\omega) = \pi(\iota_v(\iota_v T))$ where T satisfies that $\pi(T) = \omega$. Hence, we have that $\iota_v(\iota_v\omega) = 0$, and from there we can show that $\iota_{v_1}(\iota_{v_2}\omega) = -\iota_{v_2}(\iota_{v_1}\omega)$ just like before.

8/13/2025

I'm going to take a break from Guillemin's book and instead try to learn some algebraic topology. To do this I'm going to start following Munkres' Topology.

If $f_1, f_2 : X \to Y$ are continuous maps, we say f_1 is homotopic to f_2 if there is a continuous map $F : X \times [0,1] \to Y$ such that $F(x,0) = f_1(x)$ and $F(x,1) = f_2(x)$. F is called a homotopy between f_1 and f_2. And if f_1 and f_2 are homotopic, we write $f_1 \simeq f_2$. If $f_1 \simeq f_2$ and f_2 is a constant map, then we say f_1 is nulhomotopic.

An important special case is when f_1 and f_2 are paths (i.e. continuous maps from $[0,1]$ to a topological space X). In this case, it can be helpful to make the following stricter distinction. We say f_1 and f_2 are path homotopic if they have the same initial point x_0 and final point x_1, and there is a homotopy F between the two paths such that $F(0,t) = x_0$ and $F(1,t) = x_1$ for all t. Also, we call F a path homotopy and say $f_1 \simeq_p f_2$.

Lemma 51.1: $\simeq$ and $\simeq_p$ are equivalence relations.

Proof:
It's clear that any f is homotopic to itself. Also, if $F(x,t)$ is a homotopy showing that $f_1 \sim f_2$, then $G(x,t) = F(x, 1-t)$ is a homotopy showing that $f_2 \sim f_1$.

Finally, suppose $f_1 \simeq f_2$ and $f_2 \simeq f_3$. Then there exists two homotopy's $F^{(1)}$ between f_1 and f_2 and $F^{(2)}$ between f_2 and f_3. So, define:

$$G(x,t) = \begin{cases} F^{(1)}(x, 2t) & \text{for } t \in [0, 1/2] \\ F^{(2)}(x, 2t-1) & \text{for } t \in [1/2, 1] \end{cases}$$

Then G is a homotopy between f_1 and f_3, meaning $f_1 \simeq f_3$.
We know G is continuous by the pasting lemma.

The added stuff needed to $\simeq_p$ is an equivalence relation is obvious.

If f is a path in X from x_0 to x_1 and g is a path in X from x_1 to x_2, we define the product $f * g$ to be the path h given by the equation:

$$h(s) = \begin{cases} f(2s) & \text{for } s \in [0, 1/2] \\ g(2s-1) & \text{for } s \in [1/2, 1] \end{cases}$$

By the pasting lemma, h is a well-defined path in X from x_0 to x_2.

If $f : [0,1] \to X$ is a path, let $[f]$ denote the path homotopy class of f. Then the product operation induces a well-defined operation on path-homotopy classes. Specifically, given a class $[f]$ from x_0 to x_1 and a class $[g]$ from x_1 to x_2, define $[f] * [g] = [f * g]$.

To verify that this is well defined, suppose $f \simeq_p f'$ and $g \simeq_p g'$. Then if F is a homotopy from f to f' and G is a homotopy from g to g', we can define a homotopy H from $f * g$ to $f' * g'$ by the formula:

$$H(s,t) = \begin{cases} F(2s,t) & \text{for } s \in [0, 1/2] \\ G(2s-1,t) & \text{for } s \in [1/2, 1] \end{cases}$$

H is well-defined and continuous by pasting lemma.

Recall that a groupoid is a category in which every morphism is an isomorphism (look at my old Allufi notes to see what a category is...). Using the product operation of path-homotopy classes, we can define a groupoid as follows:

Consider the space X as a collection of objects, and for any $x_0, x_1 \in X$, let $\mathrm{Hom}_X(x_0, x_1)$ be the collection of path homotopy classes from x_0 to x_1. For the law of composition, say that if $[f] \in \mathrm{Hom}_X(x_0, x_1)$ and $[g] \in \mathrm{Hom}_X(x_1, x_2)$, then $[g][f] = [f * g] \in \mathrm{Hom}_X(x_0, x_2)$.

We claim:

- Every point has an identity morphism (namely the homotopy class of the constant map).
- For any $[f] \in \mathrm{Hom}(x_0, x_1)$, you can reverse the path f (i.e. define $\bar{f}(s) \coloneqq f(1-s)$) in order to get an inverse morphism in $\mathrm{Hom}(x_1, x_0)$.
- Finally, if $[f] \in \mathrm{Hom}(x_0, x_1)$, $[g] \in \mathrm{Hom}(x_1, x_2)$, and $[h] \in \mathrm{Hom}(x_2, x_3)$, then:
$$[f] * ([g] * [h]) = ([f] * [g]) * [h].$$

Proof:
We start with two lemmas:

1. If $k : X \to Y$ is a continuous map and F is a path homotopy in X between the paths f and f', then $k \circ F$ is a path homotopy in Y between the paths $k \circ f$ and $k \circ f'$.
2. If $k : X \to Y$ is a continuous map and f and g are paths in X with $f(1) = g(0)$, then $k \circ (f * g) = (k \circ f) * (k \circ g)$.

To prove the first bullet point, let $e_0 : [0,1] \to [0,1]$ be the constant function equal to 0 and $i : [0,1] \to [0,1]$ be the identity map. Then when considering both of those as paths in $[0,1]$, we can fairly easily find a path homotopy G from $e_0 * i$ to i.

One path homotopy that works is to define
$$F(s,t) = t(e_0 * i)(s) + (1-t)i(s).$$

Now suppose $e_{x_0} : [0,1] \to X$ is constant at x_0 and $f : [0,1] \to X$ is a path from x_0 to x_1. Then $e_{x_0} = f \circ e_0$, $f = f \circ i$, and by our two lemmas, $f \circ G$ is a path homotopy from $f = f \circ i$ to $f \circ (e_0 * i) = (f \circ e_0) * (f \circ i) = e_{x_0} * f$. Similar reasoning shows that if $e_{x_1} : [0,1] \to X$ is constant at x_1, then $f \simeq_p f * e_{x_1}$. This proves bullet point 1.

To prove the second bullet point, let $\bar{i} = i(1-s)$. Then we can find a homotopy G from e_0 to $i * \bar{i}$ (one that works is $G(s,t) = t((i * \bar{i})(s))$).

Then for any path f from x_0 to x_1, we can easily see that $e_{x_0} = f \circ e_0$, $f = f \circ i$, and $\bar{f} = f \circ \bar{i}$. Hence by our two lemmas, $f \circ G$ is a homotopy between $e_{x_0} = f \circ (e_0)$ and $f * \bar{f} = (f \circ i) * (f \circ \bar{i}) = f \circ (i * \bar{i})$. Also, once again similar reasoning shows that $e_{x_1} \simeq_p \bar{f} * f$. This proves bullet point 2.

I'm bored. So tldr: to prove the third bullet point just note that we can apply a continuous reparametrization $k(s)$ to $((f * g) * h)(s)$ to get $(f * (g * h))(s)$. Hence, we can define a homotopy:

$$G(s,t) \coloneqq ((f * g) * h)((1-t)s + tk(s)). \blacksquare$$

Now given a point $x_0 \in X$, define $\pi_1(X, x_0) \coloneqq \mathrm{End}(x_0)$. This is the fundamental group of X relative to x_0, and it is in fact a group with respect to our product operation since X was a groupoid. I'm going to state the next proposition as abstractly as I can cause why the hell not.

Proposition: Let **C** be a groupoid and let $A, B \in \mathrm{Obj}(\mathbf{C})$. If there exists $g \in \mathrm{Hom}(A, B)$, then $\mathrm{End}(A) \cong \mathrm{End}(B)$.

Proof:
If $f \in \mathrm{End}(A)$, then define $\phi(f) = gfg^{-1}$. Then it's clear that ϕ is a group homomorphism from $\mathrm{End}(A)$ to $\mathrm{End}(B)$. To show that ϕ is injective, suppose $\phi(f) = e_B$ where e_B is the identity morphism on B. Then $f = g^{-1}e_B g = g^{-1}g = e_A$ where e_A is the identity morphism on A. Next, to show that ϕ is surjective, suppose $h \in \mathrm{End}(B)$. Then $f \coloneqq g^{-1}hg$ satisfies that $\phi(f) = h$.

Corollary: If X is path connected, then $\pi_1(X, x_0) \cong \pi_1(X, x_1)$ for all $x_0, x_1 \in X$.

We say a space X is simply connected if X is path connected and for some $x_0 \in X$, $\pi_1(X, x_0) = \{1\}$.

Lemma 52.3: Let X be a path-connected topological space. Then X is simply connected iff every pair of paths f and f' with the same initial and final point are path homotopic.

$(\Longrightarrow)$
Suppose f and f' are paths from x_0 to x_1. Then if we set $\bar{f}$ and $\bar{f}'$ to be the reversed paths, we know that $[f' * \bar{f}], [f * \bar{f}'] \in \pi_1(X, x_0)$. But now since $\pi_1(X, x_0)$ is trivial, we know that $[f' * \bar{f}] = 1 = [f * \bar{f}']$. So, there is a path homotopy F between $f' * \bar{f}$ and $f * \bar{f}$. Now just define $G(s,t) = F(\frac{1}{2}s, t)$ and we have shown that f and f' are path homotopic.

$(\Longleftarrow)$
Suppose $f \in \pi_1(X, x_0)$ for some $x_0 \in X$. Also suppose g is a path in X from x_0 to x_1 where $x_1 \neq x_0$. (Note, this lemma is trivial if X has only one point. So, we can without loss of generality assume X has more than one point.)

Then since both g and $f * g$ are paths from x_0 to x_1, we know there is a homotopy F between the two paths. So, $[g] = [f * g] = [g] * [f]$. If we apply on the left side the class $[\bar{g}]$, then this means that $1 = [f]$. Hence, $\pi_1(X, x_0)$ is trivial. ■

A consequence of this lemma is that all convex subsets of $\mathbb{R}^n$ are simply connected.

Recall the two lemmas stated on page 118. Those lemmas will let us define an important thing.

Suppose $h : X \to Y$ is a continuous map such that $h(x_0) = y_0$. We will denote this by writing $h : (X, x_0) \to (Y, y_0)$. Then we define the homomorphism induced by h, $h_* : \pi_1(X, x_0) \to \pi_2(X, y_0)$, by the map:

$$h_*([f]) \coloneqq [h \circ f]$$

This is well defined since if F is a homotopy between f and f', then $h \circ F$ is a homotopy between $h \circ f$ and $h \circ f'$. Also, this is indeed a group homomorphism since:

$$h_*([f] * [g]) = [h \circ (f * g)] = [h \circ f] * [h \circ g] = h_*([f]) * h_*([g])$$

Theorem 52.4: If $h : (X, x_0) \to (Y, y_0)$ and $k : (Y, y_0) \to (Z, z_0)$ are continuous maps, then $(k \circ h)_* = k_* \circ h_*$. Also if $i : (X, x_0) \to (X, x_0)$ is the identity map on X, then i_* is the identity map on $\pi_1(X, x_0)$.

Hopefully this is self-explanatory.

Corollary 52.5: If $h : (X, x_0) \to (Y, y_0)$ is a homeomorphism from X to Y, then h_* is an isomorphism from $\pi_1(X, x_0)$ to $\pi_1(Y, y_0)$.

Proof:
Let $k = h^{-1}$. Then from the last theorem we can easily see that k_* and h_* are inverses of each other. Hence, h_* is invertible and thus a group isomorphism. ■

Consequently, we know that the fundamental group of a path connected space is a topological property. So, if two path connected spaces do not have the same fundamental group, they can't be homeomorphic.

8/15/2025

If α is a path from x_0 to x_1 in X, we shall denote $\bar{\alpha}$ to be the reversed path from x_1 to x_0. Also we shall denote $\hat{\alpha} : \pi_1(X, x_0) \to \pi_1(X, x_1)$ to be the isomorphism $[f] \mapsto [\bar{\alpha} * f * \alpha]$.

Exercise 52.3: Let $x_0 \to x_1$ be points of the path-connected space X. Show that $\pi_1(X, x_0)$ is abelian iff for every pair α and β of paths from x_0 to x_1, we have $\hat{\alpha} = \hat{\beta}$.

$(\Longrightarrow)$
Suppose $\pi_1(X, x_0)$ is abelian and we have two paths α, β from x_0 to x_1. Then for any $[f] \in \pi_1(X, x_0)$, we have that $[f] * [\alpha * \bar{\beta}] = [\alpha * \bar{b}] * [f]$. Therefore:

$$\hat{\alpha}(f) = [\bar{\alpha} * f * \alpha] = [\bar{\beta} * f * \beta] = \hat{\beta}(f).$$

$(\Longleftarrow)$
Suppose $[f] \in \pi_1(X, x_0)$ and α, β are paths from x_0 to x_1. Then since $\hat{\alpha}([f * \alpha * \bar{\beta}]) = \hat{\beta}([f * \alpha * \bar{\beta}])$, we have that:

$$[\bar{\alpha} * f * \alpha * \bar{\beta} * \alpha] = [\bar{\beta} * f * \alpha * \bar{\beta} * \beta] = [\bar{\beta} * f * \alpha]$$

Hence $[f] * [\alpha * \bar{\beta}] = [\alpha * \bar{\beta}] * [f]$, and this proves that the group operation of $\pi_1(X, x_0)$ is commutative so long as one of the arguments passes through x_1.

Now to prove general commutativity, suppose $[f], [g] \in \pi_1(X, x_0)$ and α is a path from x_0 to x_1. Then from before we know that:

$$[f] * [g] = [f] * [g * \alpha * \bar{\alpha}] = [g * \alpha * \bar{\alpha}] * [f] = [g] * [f].$$

Exercise 52.4: Let $A \subseteq X$, suppose $r : X \to A$ is a continuous map such that $r(a) = a$ for each $a \in A$. (The map r is called a retraction of X onto A and we call A a retract of X.) If $a_0 \in A$, then show that $r_* : \pi_1(X, a_0) \to \pi_1(A, a_0)$ is surjective.

Suppose $[f]_A \in \pi_1(A, a_0)$. Then f is also a loop in X, meaning it has a class $[f]_X \in \pi_1(X, a_0)$. And $r_*([f]_X) = [r \circ f]_A = [f]_A$. Hence r_* is surjective.

Exercise 52.5: Let A be a subspace of a simply connected space X, and let $h : (A, a_0) \to (Y, y_0)$ be a continuous map. If h is extendable to a continuous map of X into Y, then h_* is the trivial homomorphism (i.e. the homomorphism mapping everything to the identity element).

Let h' be a continuous extension of h to all of X. Now importantly, for any loop $[f]_A \in \pi_1(A, a_0)$, we know f also defines a class $[f]_X \in \pi_1(X, a_0)$. Also importantly:

$$h'_*([f]_X) = [h' \circ f]_Y = [h \circ f]_Y = h_*([f]_A).$$

However, since X is simply connected, we know $[f]_X = 1$ and thus $h'_*([f]_X) = 1$. So, we've shown that $h_*([f]_A) = 1$ for all $[f]_A \in \pi_1(A, a_0)$.

Note: Munkres specifically sets $X = \mathbb{R}^n$ in his statement of the exercise.

Exercise 52.6: Suppose $h : X \to Y$ is a continuous map, α is a path from x_0 to x_1 in X, and $\beta := h \circ \alpha$. Then $h_* \circ \hat{\alpha} = \hat{\beta} \circ h_*$.

Suppose $[f] \in \pi_1(X, x_0)$. Then:

$$\begin{aligned} h_*(\hat{\alpha}([f])) = [h \circ (\bar{\alpha} * f * \alpha)] &= [(h \circ \bar{\alpha}) * (h \circ f) * (h \circ \alpha)] \\ &= [\bar{\beta} * h_*([f]) * \beta] = \hat{\beta}(h_*([f])). \end{aligned}$$

Let $p : E \to B$ be a continuous surjective map. Then an open set $U \subseteq B$ is said to be evenly covered by p if $p^{-1}(U)$ is a union of disjoint open sets $V_\alpha \subseteq E$ satisfying that $p|_{V_\alpha}$ is a homeomorphism from V_α to U. The collection $\{V_\alpha\}_{\alpha \in A}$ will be called a partition of $p^{-1}(U)$ into slices.

Note that if $W \subseteq U$ is also open, then W is also openly covered by p with there being a partition of slices $\{V_\alpha \cap p^{-1}(W)\}_{\alpha \in A}$.

If every point $b \in B$ has an open neighborhood $U \subseteq B$ that is evenly covered by p, we call p a covering map and E a covering space of B.

Claim: If $p : E \to B$ is a covering map of B, then p is an open map.

Proof:

Suppose $A \subseteq E$ is open. Then for any $y \in f(A)$, there exists an open neighborhood U of y that is evenly covered by p. In turn there exists $x \in A$ with $p(x) = y$ and an open neighborhood V_α of x such that $p|_{V_\alpha}$ is a homeomorphism from V_α to U_y. And hence, $p(A \cap V_\alpha)$ is an open neighborhood of y in $U \subseteq B$ that is also a subset of $p(A)$. It follows that $p(A)$ is an open subset of B. ■

Corollary: If $p : E \to B$ is a covering map of B, then p is a local homeomorphism, meaning each point $e \in E$ has an open neighborhood that is mapped homeomorphically by p onto an open subset of B.

Proof:

If $x \in E$, then the reasoning from the prior proof lets us pick an open set $A \cap V_\alpha$ containing x such that, $p|_{A \cap V_\alpha}$ is a homeomorphism to an open set $p(A \cap V_\alpha)$ in B.

Theorem 53.1: The map $p : \mathbb{R} \to S^1$ given by $p(x) = (\cos(2\pi x), \sin(2\pi x))$ is a covering map of S^1.

Hopefully this is obvious. We can cover S^1 with four open sets gotten by intersecting S^1 with the top, bottom, right, and left open halfs respectively of the coordinate plane. Then it's easy to find a partition of each preimage into slices.

As a side note, if $H^1 = \{x \in \mathbb{R} : x \geq 0\}$, then $p|_{H^1}$ is surjective and a local homeomorphism. That said, it is not a covering map since the point $(1, 0) \in S^1$ has no open neighborhood U that is evenly covered by $p|_{H^1}$.

(The specific issue we run into is that for any open set U we pick, $p|_{H^1}$ will not be a surjective map from the slice of the preimage containing 0 to U…)

Theorem 53.2: Let $p : E \to B$ be a covering map. If B_0 is a subspace of B and $E_0 := p^{-1}(B_0)$, then the map $p_0 : E_0 \to B_0$ obtained by restricting p is a covering map.

Proof:

Given $y \in B_0$, let $U \subseteq B$ be an open neighborhood of y that is evenly covered and let $\{V_\alpha\}_{\alpha \in A}$ be a partition of $p^{-1}(U)$ into slices. Then $U \cap B_0$ is an open neighborhood of y in B_0 that is evenly covered by p_0 via the partition $\{V_\alpha \cap E_0\}_{\alpha \in A}$ of $p^{-1}(U \cap B_0)$ into slices. ■

Theorem 53.3: If $p : E \to B$ are covering maps and $p' : E' \to B'$ are covering maps, then the map $p \times p' : E \times E' \to B \times B'$ defined by $(e, e') \mapsto (p(e), p'(e'))$ is a covering map.

Proof:

Given $b \in B$ and $b' \in B'$, let U and U' be open neighborhoods of b and b' that are evenly covered by p and p' respectively. Next let $\{V_\alpha\}_{\alpha \in A}$ and $\{V'_\gamma\}_{\gamma \in C}$ be partitions of $p^{-1}(U)$ and $(p')^{-1}(U')$ respectively into slices. Then:

$$(p \times p')^{-1}(U \times U') = \bigcup_{\alpha \in A} \Big(\bigcup_{\gamma \in C} (V_\alpha \times V'_\gamma) \Big).$$

Also $U \times U'$ is open in $B \times B'$; $V_\alpha \times V'_\gamma$ is open in $E \times E'$ for all α and γ; the $V_\alpha \times V'_\gamma$ are all disjoint; and each $V_\alpha \times V'_\gamma$ is mapped homeomorphically onto $U \times U'$. ■

Exercise 53.2: Let $p : E \to B$ be continuous and surjective, and suppose $U \subseteq B$ is an open set that is evenly covered by p. If U is connected, then the partition of $p^{-1}(U)$ into slices is unique.

Proof:

For the sake of contradiction, suppose $\{V_\alpha\}_{\alpha \in A}$ and $\{V'_\beta\}_{\beta \in B}$ are two different partitions of $p^{-1}(U)$ into slices. Then we know there exists $V_{\alpha_0}, V'_{\beta_0}$ in those two partitions satisfying that $V_{\alpha_0} \cap V'_{\beta_0} \neq \emptyset$ and $V_{\alpha_0} \neq V'_{\beta_0}$. Without loss of generality suppose $V_{\alpha_0} - V'_{\beta_0} \neq \emptyset$. Then since p is a homeomorphism from V_{α_0} to U, we know $p(V_{\alpha_0} \cap V'_{\beta_0})$ is open in U. Also, since $V'_{\beta_0} = p^{-1}(U) - (\bigcup_{\beta \neq \beta_0} V'_\beta)$, we know V'_{β_0} is closed in $p^{-1}(U)$. Hence $V_{\alpha_0} - V'_{\beta_0}$ is open in V_{α_0} and so $p(V_{\alpha_0} - V'_{\beta_0})$ is also open in U.

But now $p(V_{\alpha_0} - V'_{\beta_0})$ and $p(V_{\alpha_0} \cap V'_{\beta_0})$ are two disjoint nonempty open subsets of U. This contradicts that U is connected. ■

Exercise 53.3: Let $p : E \to B$ be a covering map and let B be connected. If $p^{-1}(\{b_0\})$ has k elements for some $b_0 \in B$, then $p^{-1}(\{b\})$ has k elements for all $b \in B$. In such a case E is called a k-fold covering of B.

Proof:

Let $B_k := \{b \in B : |p^{-1}(\{b\})| = k\}$. Now it's very clear that if $b \in B$ and $U \subseteq B$ is an open neighborhood of b that is evenly covered by p, then $U \subseteq B_k$. Hence, B_k is open. Also note that if $b \in (B_k)^\mathsf{C}$ and $U \subseteq B$ is an open neighborhood of b that is evenly covered by p, then $U \subseteq (B_k)^\mathsf{C}$. Hence $(B_k)^\mathsf{C}$ is open. Since B is connected, this means that B_k or $(B_k)^\mathsf{C}$ must be empty. and since $b_0 \in B_k$, we know the empty set isn't B_k. ■

Let $p : E \to B$ be a map. If $f : X \to B$ is a continuous map, a lifting of f is a map $\tilde{f} : X \to E$ such that $p \circ \tilde{f} = f$. Or in other words, the following diagram commutes:

$$\begin{tikzcd} & E \arrow[d, "p"] \\ X \arrow[ur, "\tilde{f}"] \arrow[r, "f"'] & B \end{tikzcd}$$

Lemma 54.1: Let $p : E \to B$ be a covering map and let $e_0 \in p^{-1}(\{b_0\})$. Then if $f : [0,1] \to B$ is a path beginning at b_0, there is a unique lifting of f to a path $\tilde{f} \in E$ beginning at e_0.

Proof:

Let $\{U_\alpha\}_{\alpha \in A}$ be an open covering of B consisting of open sets that are evenly covered by p. Then in turn $\{f^{-1}(U_\alpha)\}_{\alpha \in A}$ is an open covering of $[0,1]$. So, we can invoke the following useful lemma.

<u>Lebesgue Number Lemma:</u> If (X,d) is a compact metric space and an open cover $\mathcal{U}$ of X is given, then the cover admits some <u>Lebesgue number</u> $\delta > 0$. That is, for some $\delta > 0$ we have that: $\mathrm{diam}(E) < \delta \Longrightarrow \exists U \in \mathcal{U}\ s.t.\ E \subseteq U$.

Proof:

Let $\{U_1, \ldots, U_n\} \subseteq \mathcal{U}$ be a finite subcover of X. If any $U_i = X$, then trivially any $\delta > 0$ works. So assume $U_i \neq X$ for all i. Then since $F_i := X - U_i$ is nonempty for all i, we know that $d(x, F_i) = \inf\{d(x,y) : y \in F_i\}$ is well defined for all $x \in X$ and $i \in \{1, \ldots, n\}$. Furthermore, it is easy to see that $d(x, F_i)$ is a continuous function of x. And since F_i is closed, we have that $d(x, F_i) = 0$ iff $x \in F_i$.

Now define $f : X \to \mathbb{R}$ by $f(x) = \sum\limits_{i=1}^{n} d(x, F_i)$.

Since f is a continuous map from a compact set, we know by the extreme value theorem that f attains a minimum $\alpha \in \mathbb{R}$. Also, since the F_i cover X, we know that $f(x) \neq 0$ for any $x \in X$. Hence, $\alpha > 0$.

Now set $\delta = {}^{\alpha}/_{n}$. Then for any $E \subseteq X$ with $\mathrm{diam}(E) < \delta$, if we pick $x_0 \in E$ we have that $E \subseteq B_\delta(x_0)$. Importantly, since $f(x_0) > \alpha = n\delta$, we must have that $d(x, F_i) > \delta$ for some i. It follows then that $E \subseteq U_i$. This proves that δ works as a Lebesgue number. ■

By the above lemma, we know that there exists $0 = s_0 < s_1 < \ldots < s_n = 1$ satisfying for each $0 \leq i < n$ that $[e_i, e_{i+1}] \subseteq f^{-1}(U_{\alpha_i})$ for some $\alpha_i \in A$. Or in other words, $f([e_i, e_{i+1}]) \subseteq U_{\alpha_i}$ for some $\alpha_i \in A$. After setting $\tilde{f}(0) = e_0$, we may proceed by induction as follows in order to show a suitable $\tilde{f}$ exists.

Assume that $\tilde{f}$ was already defined for all $s \in [0, s_i]$ where $i < n$. Since U_{α_i} is evenly covered by p, we know there is a unique open set $V \subseteq E$ such that $\tilde{f}(s_i) \in V$ and $p|_V$ is a homeomorphism onto U_{α_i}. Then since $f([s_i, s_{i+1}]) \subseteq U_{\alpha_i}$, we know it is well defined to set $\tilde{f}(s) = (p|_V)^{-1}(f(s))$ for all $s \in [s_i, s_{i+1}]$. Also, it is clear then that $\tilde{f}|_{[s_i, s_{i+1}]}$ is continuous since it is the composition of two continuous functions. In turn, we can see by the pasting lemma that $\tilde{f}$ will be continuous on the domain $[0, s_{i+1}]$

To finish off, we now need to show that $\tilde{f}$ is unique. So suppose $g : [0,1] \to E$ also satisfies that $p \circ g = E$ and $g(0) = e_0$. Then we trivially have that $g(s_0) = \tilde{f}(s_0)$. So, we may proceed by induction as follows in order to show $\tilde{f} = g$.

Assume we've already shown that $\tilde{f}(s) = g(s)$ for all $s \in [0, s_i]$ where $0 \leq i < n$. Now letting V be as in the prior reasoning, we know that $g(s)$ must be in V for all $s \in [s_i, s_{i+1}]$. After all, if $\{V_\gamma\}_{\gamma \in C}$ is a partition of $p^{-1}(U_{\alpha_i})$ into slices, we must have that:

$$g([s_i, s_{i+1}]) \subseteq p^{-1}(U_{\alpha_i}) = \bigcup\nolimits_{\gamma \in C} V_\gamma.$$

But since all the V_γ are disjoint, nonempty, and open; and $g([s_i, s_{i+1}])$ is connected, it must be the case that $g([s_i, s_{i+1}])$ only intercepts one V_γ. Specifically, that one V_γ is V since $g(s_i) \in V$.

But now we must have for each $s \in [s_i, s_{i+1}]$ that $g(s)$ satisfies that $p(g(s)) = f(s)$. Yet the only point in V which satisfies that is $\tilde{f}(s)$. Hence, $\tilde{f}(s) = g(s)$ for all $[0, s_{i+1}]$. ■

Lemma 54.2: Let $p : E \to B$ be a covering map and let $e_0 \in p^{-1}(\{b_0\})$. Then if $F : [0,1]^2 \to B$ is a continuous map satisfying that $F(0,0) = b_0$, there is a unique lifting of F to a continuous function $\tilde{F} : [0,1]^2 \to E$ such that $\tilde{F}(0,0) = e_0$.

Proof:
We'll start by showing uniqueness. Suppose $\tilde{F}$ and G both are continuous functions from $[0,1]^2$ to E satisfying our lemma. By an easy application of the last lemma, if $(x_0, y_0) \in [0,1]^2$, then we must have $\tilde{f}(s) := \tilde{F}(sx_0, sy_0)$ and $g(s) := G(sx_0, sy_0)$ are equal for all $s \in [0,1]$ since both are the unique lifting of the path $f(s) := F(sx_0, sy_0)$ to a path in E starting at e_0. But then we've shown that $\tilde{F}(x_0, y_0) = \tilde{f}(1) = g(1) = G(x_0, y_0)$. And since (x_0, y_0) was arbitrary, we have shown that $\tilde{F}$ is unique if it exists.

Next, we need to show that a sufficient $\tilde{F}$ exists in the first place. So start by setting $\tilde{F}(0,0) = e_0$. Now for any fixed $(x_0, y_0) \in [0,1]^2$, if we define $f(s) = F(sx_0, sy_0)$, then we know by the prior lemma that there is a continuous map $\tilde{f} : [0,1] \to E$ such that $p \circ \tilde{f}(s) = f(s)$ for all $s \in [0,1]$. Hence, we may define $\tilde{F}(x_0, y_0) = \tilde{f}(1)$. After doing this for all choices of (x_0, y_0), we will have constructed a function $\tilde{F} : [0,1]^2 \to E$ such that $p \circ \tilde{F} = F$.

What's still not clear is that $\tilde{F}$ is continuous. To show this, we will need the observation that if $f(s) = F(sx_0, sy_0)$ and $\tilde{f}$ is the unique lifting of f to path in E such that $p \circ \tilde{f} = f$ and $\tilde{f}(0) = e_0$, then by an easy application of the last lemma we can show that $\tilde{F}(sx_0, sy_0) = \tilde{f}(s)$. Hence, for any $(x_0, y_0) \in [0,1]^2$ we have that $\tilde{F}$ varies continuously as one moves along the straight line between $(0,0)$ to (x_0, y_0).

Now use the Lebesgue number lemma to pick $0 = s_0 < s_1 < \ldots < s_n = 1$ such that for all $0 \leq i, j < n$, $F([s_i, s_{i+1}] \times [s_j, s_{j+1}])$ is contained in some open set U that is evenly covered by p. Then make the inductive hypotheses that there exists $0 \leq i, j < n$ for which we've already shown that $\tilde{F}$ is continuous on:

$$([0, s_i] \times [0,1]) \cup ([s_i, s_{i+1}] \times [0, s_j]).$$

Note that if $i = 0$ or $j = 0$, then the fact that $\tilde{F}$ is continuous on the line connecting $(0,0)$ to $(0,1)$ and on the line connecting $(0,0)$ to $(1,0)$ proves this base case.

Now let U be an open set that contains $F([s_i, s_{i+1}] \times [s_j, s_{j+1}])$ and is evenly covered by p. Then let $V \subseteq E$ be an open set disjoint from the rest of the preimage of U such that $p|_V$ is a homeomorphism from V to U and $\tilde{F}(s_i, s_j) \in V$. Since we already know by induction that F varies continuously along the straight lines from (s_i, s_j)

to (s_i, s_{j+1}), and from (s_i, s_j) to (s_{i+1}, s_j), we know that $F(x, s_j)$ and $F(s_i, y)$ are in V for all $x \in [s_i, s_{i+1}]$ and $y \in [s_j, s_{j+1}]$. This is important because we know that for any $(x, y) \in [s_i, s_{i+1}] \times [s_j, s_{j+1}]$, the straight line from $(0, 0)$ to (x, y) must cross one of those two borders of the rectangle. And since F varies continuously along the straight line from $(0, 0)$ to (x, y), this proves that $F(x, y) \in V$ for all $(x, y) \in [s_i, s_{i+1}] \times [s_j, s_{j+1}]$.

In turn, we have for all $(x, y) \in [s_i, s_{i+1}] \times [s_j, s_{j+1}]$ that $\tilde{F} = (p|_V)^{-1}(F(x, y))$. Thus $\tilde{F}$ is continuous on $[s_i, s_{i+1}] \times [s_j, s_{j+1}]$ since it is the composition of two continuous functions. Also by pasting lemma, we can thus say that $\tilde{F}$ is continuous on $([0, s_i] \times [0, 1]) \cup ([s_i, s_{i+1}] \times [0, s_{j+1}])$. ■

Corollary: If F is a path homotopy then the lifting in the prior lemma: $\tilde{F}$, is a path homotopy.

Proof:
If $F(0, t) = b_0$ for all $t \in [0, 1]$, then we know that $\tilde{F}(\{0\} \times [0, 1]) \subseteq p^{-1}(b_0)$. But the latter set will have the discrete topology as a subspace of E. Since $\tilde{F}(\{0\} \times [0, 1])$ is connected, this must mean that $\tilde{F}$ is constant on $0 \times [0, 1]$. Similar reasoning also shows that $\tilde{F}(\{1\} \times [0, 1])$ has one element. ■

Note: $p^{-1}(b_1)$ as a subspace of E has the discrete topology because p being a covering map means that each of the elements of $p^{-1}(b_1)$ are contained in distinct disjoint open sets. Also, a subset of a space with the discrete topology is connected iff it has one element.

Theorem 54.3: Let $p : E \to B$ be a covering map, and let $e_0 \in p^{-1}(b_0)$. Next let f and g be paths in B from b_0 to b_1 and let $\tilde{f}$ and $\tilde{g}$ be their respective liftings to a path in E starting at e_0. If f and g are path homotopic, then $\tilde{f}$ and $\tilde{g}$ end at the same point and are path homotopic.

Proof:
Let $F : [0, 1]^2 \to B$ be a homotopy between f and g. Then $F(0, 0) = b_0$, meaning there is a unique continuous lifting $\tilde{F} : [0, 1]^2 \to E$ of F to E satisfying that $\tilde{F}(0, 0) = e_0$. By our prior corollary, we know that $\tilde{F}$ will be a homotopy, meaning that $\tilde{f}(\{0\} \times [0, 1]) = \{e_0\}$ and $\tilde{f}(\{1\} \times [0, 1]) = \{e_1\}$ where $e_1 \in p^{-1}(b_1)$. Also, due to the uniqueness we proved in lemma 54.1, it's easy to see that $\tilde{F}(s, 0) = \tilde{f}(s)$ and $\tilde{F}(s, 1) = \tilde{g}(s)$ for all s. ■

Let $p : E \to B$ be a covering map and let $b_0 \in B$. Choose an $e_0 \in E$ such that $p(b_0) = b_0$. Then given an element $[f] \in \pi_1(B, b_0)$, define $\phi([f]) := \tilde{f}(1)$ where $\tilde{f}$ is the unique lifting of f to path in E starting at e_0.

By the last theorem, $\phi : \pi_1(B, b_0) \to p^{-1}(b_0)$ is a well-defined map which we call the lifting correspondence derived from p. Note, ϕ is dependent on our choice of e_0.

Theorem 54.4: Let $p : E \to B$ be a covering map and for some $b_0 \in B$, choose some $e_0 \in p^{-1}(b_0)$. If E is path connected, then the lifting correspondence $\phi : \pi_1(B, b_0) \to p^{-1}(b_0)$ is surjective. Furthermore, if E is simply connected, then ϕ is bijective.

Proof:
If E is path connected, then given any $e_1 \in p^{-1}(b_0)$ there exists a path $\tilde{f}$ from e_0 to e_1. In turn, $f := p \circ \tilde{f}$ is a continuous loop from b_0 to itself in B such that $\phi([f]) = \tilde{f}(1) = e_1$.

Next suppose E is simply connected. Then suppose $[f], [g] \in \pi_1(B, b_0)$ satisfy that $\phi([f]) = \phi([g])$. By letting $\tilde{f}$ and $\tilde{g}$ be the liftings of f and g respectively, we know that $\tilde{f}$ and $\tilde{g}$ are both paths from e_0 to some e_1. Therefore, by lemma 52.3 plus the fact that E is simply connected, we know that $\tilde{f}$ and $\tilde{g}$ are path homotopic via a homotopy $\tilde{F}$. And since p is a continuous map, we have that $F := p \circ \tilde{F}$ is a path homotopy of $f = p \circ \tilde{f}$ and $g = p \circ \tilde{g}$. This proves that $[f] = [g]$. ■

Theorem 54.5: The fundamental group of $S^1 := \{x \in \mathbb{R}^2 : \|x\|_2 = 1\}$ is isomorphic to $\mathbb{Z}$.

Proof:
Let $p : \mathbb{R} \to S^1$ be the covering map $p(t) = (\cos(2\pi t), \sin(2\pi t))$. Also pick $e_0 = 0$ and let $b_0 = p(e_0) = (1, 0)$. Then $p^1(\{b_0\}) = \mathbb{Z}$. And since $\mathbb{R}$ is simply connected, we have that the lifting correspondence $\phi : \pi_1(S^1, b_0) \to \mathbb{Z}$ is bijective.

To show that $\pi_1(S^1, b_0) \cong \mathbb{Z}$, we shall now show that ϕ is a group homomorphism.

Suppose $[f], [g] \in \pi_1(S^1, b_0)$. Then let $\tilde{f}$ and $\tilde{g}$ be the respective liftings of f and g to $\mathbb{R}$ to paths starting at 0. Also let $n := \phi([f])$ and $m := \phi([g])$. If we define $\tilde{h}(s) = \tilde{g}(s) + n$, then because $p(x + n) = p(x)$ for all $x \in \mathbb{R}$, we know $\tilde{h}$ is another lifting of g. However $\tilde{h}$ starts at n instead of 0. It follows that $\tilde{f} * \tilde{h}$ is a well-defined product of paths, and it is lifting of $f * g$ to a path in $\mathbb{R}$ starting at 0. Hence:

$$\phi([f] * [g]) = \phi([f * g]) = (\tilde{f} * \tilde{h})(1) = n + m = \phi([f]) + \phi([g]). \blacksquare$$

The intution for this result is that the fundamental group of S^1 categorizes all loops in S^1 by the net number of complete revolutions of the path around S^1.

Theorem 54.6: Let $p : E \to B$ be a covering map and let $p(e_0) = b_0$.

(a) The induced homomorphism $p_* : \pi_1(E, e_0) \to \pi_1(B, b_0)$ is injective.

Proof:
Suppose $[\tilde{h}] \in \pi_1(E, e_0)$ such that $p_*([\tilde{h}])$ is the identity element of $\pi_1(B, b_0)$. Then there is a homotopy F between $p \circ \tilde{h}$ and the constant loop based at b_0. If $\tilde{F}$ is the lifting of F to E satisfying that $\tilde{F}(0, 0) = e_0$, then $\tilde{F}$ will be a homotopy between $\tilde{h}$ and the constant loop based at e_0.

(b) Let $H = p_*(\pi_1(E, e_0))$. Then if $\pi_1(B, b_0)/H$ is collection of right cosets of H, then the lifting correspondance of ϕ induces an injective map:

$$\Phi : \pi_1(B, b_0)/H \to p^{-1}(b_0)$$

Furthermore, Φ is bijective, if E is path connected.

Proof:
Let f and g be loops in B based at b_0, and let $\tilde{f}$ and $\tilde{g}$ be the liftings of those loops in E to paths starting at e_0. Since H is a subgroup of $\pi_1(B, b_0)$, we know that $[g] \in H * [f]$ iff $H * [g] = H * [f]$. Thus, to show that Φ is well-defined, we just need to show that $[g] \in H * [f] \Longrightarrow \phi([g]) = \phi([f])$. Meanwhile, to show that Φ is injective, we just need to show that $[g] \in H * [f] \Longleftarrow \phi([g]) = \phi([f])$.

$(\Longrightarrow)$
If $[f] \in H * [g]$, then there exists $[h] \in H$ with $[f] = [h * g]$. And due to how we defined H, the lifting $\tilde{h}$ of h to E starting e_0 will be a loop.

Why: We know there exists a loop $\tilde{h}'$ in E based at e_0 such that $[h] = [p \circ \tilde{h}']$. But now by theorem 54.3, we have that the liftings $\tilde{h}$ and $\tilde{h}'$ of h and $p \circ \tilde{h}'$ respectively are path homotopic. So, $\tilde{h}$ is also a loop based at e_0.

It follows that the product $\tilde{h} * \tilde{g}$ is well-defined and is a lifting of $h * g$. That plus the fact that $[f] = [h * g]$ means that by theorem 54.3, $\tilde{f}$ and $\tilde{h} * \tilde{g}$ have the same end point. So:

$$\phi([f]) = \tilde{f}(1) = \tilde{h} * \tilde{g}(1) = \tilde{g}(1) = \phi([g])$$

$(\Longleftarrow)$
Next suppose $\phi([f]) = \phi([g])$. Then $\tilde{f}$ and $\tilde{g}$ end at the same point of E. So, we can define the loop $\tilde{h}$ in E based at e_0 as the product of $\tilde{f}$ and the reverse of $\tilde{g}$. Importantly, $[\tilde{h} * \tilde{g}] = [\tilde{f}]$. So, if $\tilde{F}$ is a path homotopy between $\tilde{h} * \tilde{g}$ and $\tilde{f}$, then $F := p \circ \tilde{F}$ is a path homotopy between $f = p \circ \tilde{f}$ and $p \circ (\tilde{h} * \tilde{g}) = (p \circ \tilde{h}) * g$. Also, $p \circ \tilde{h} \in H$. So $[f] \in H * [g]$.

If E is path connected, then we know ϕ is surjective. In turn Φ will also be surjective.

(c) If f is a loop in B based at b_0, then $[f] \in H$ iff f lifts to a loop in E based at e_0.

Proof:
Since Φ is injective and the constant loop about b_0 is mapped to e_0 by ϕ, we know that $\phi([f]) = e_0$ iff $[f] \in H$. But $\phi([f]) = e_0$ precisely iff f lifts to a loop in E based at e_0. ■

8/17/2025

Lemma 55.1: If A is a retract of X, then the homomorphism of fundamental groups induced by the inclusion map $j : A \to X$ is injective.

Proof:
If $r : X \to A$ is a retraction, then $r \circ j$ is just the identiy map on A. It follows that $(r \circ j)_* = r_* \circ j_*$ is the identiy map on $\pi_1(A, a_0)$ for any $a_0 \in A$. This is only possible if r_* is surjective (which we admittedly already proved on page 121) and j_* is injective. ■

We'll denote $D^2 := \{x \in \mathbb{R}^2 : \|x\|_2 \leq 1\}$.

Theorem 55.2: There is no retraction from D^2 to S^1.

Proof:
If such a retraction did exist, then we would know by the previous lemma that there exists an injective group homomorphism from the fundamental group of S^1 to the fundamental group of D^2. However, the fundamental group of S^1 has strictly greater cardinality than that of D^2. So, no such injection exists. ■

Lemma 55.3: Suppose $h : S^1 \to X$ is a continuous map. Then the following are equivalent.

(1) h is nulhomotopic (meaning there is a homotopy from h to a constant function on X).

(2) h extends to a continuous map $k : D^2 \to X$.

(3) h_* is the trivial homomorphism of fundamental groups.

$(1 \implies 2)$
Let $H : S^1 \times [0,1] \to X$ be a homotopy from h to some constant map. Then if we define $\pi : S^1 \times [0,1] \to D^2$ by $\pi(x,t) = x(1-t)$, we have that π defines a quotient map.

It's obvious that π is continuous and surjective. Also, it's clear that π is an open function since any closed set of $S^1 \times [0,1]$ is compact and thus maps to another compact set of D^2 which must be closed.

Now since H is constant over the preimage with respect to π of any given singleton, we know that H induces a continuous map $k : D^2 \to X$ satisfying that $H = k \circ \pi$. Also, it's clear that for $x \in S^1$, $k(x) = H(x,0) = h(x)$.

$(2 \implies 3)$
Let $k : D^2 \to X$ be an extension of h and let $j : S^1 \to D^2$ be the inclusion map. Then $h = k \circ j$, meaning $h_* = k_* \circ j_*$. But now since the fundamental group of D^2 is trivial, we know that h_* must be the trivial homomorphism.

$(3 \implies 1)$
Let $p : \mathbb{R} \to S^1$ be the standard covering map (the one defined in theorem 53.1). Then $p_0 := p|_{[0,1]}$ is a generator for $\pi_1(S^1, 0)$. If we let $x_0 = h(b_0)$, then $f = h \circ p_0$ is a loop in X. And since h_* is trivial, we know there exists a path homotopy F from $h \circ p_0$ to the constant map x_0.

Now the map $(p_0 \times i) : [0,1]^2 \to S^1 \times [0,1]$ defined by $(p_0 \times i)(x,t) = (p_0(x), t)$ is a quotient map.

It's continuous and surjective. Also, once again it maps any closed set to another closed set since all closed subsets of $[0,1]^2$ are compact and $S^1 \times [0,1]$ is Hausdorff.

Since F is constant on $(p_0 \times i)^{-1}(y)$ for any $y \in S^1 \times [0,1]$, we have that F induces a continuous map $H : S^1 \times [0,1] \to X$ satisfying that $H \circ (p_0 \times i) = F$. Also, $x_0 = F(s,1) = H(p_0(s), 1)$ for all $s \in [0,1]$. This shows that H is our desired homotopy. ■

Corollary 55.4: The inclusion map $j : S^1 \to \mathbb{R} - \{0\}$ is not nulhomotopic. Also the identity map $i : S^1 \to S^1$ is not nulhomotopic.

Proof:
There is a retraction $k : \mathbb{R} - \{0\} \to S^1$ given by $k(x) = x/\|x\|_2$. It follows by lemma 55.1 that j_* is injective. And since the fundamental group of S^1 is not trivial, that means that j_* is not the trivial homomorphism. Hence, j isn't nulhomotopic by the last lemma.

Similarly, i_* is the identity homomorphism and thus isn't trivial.So i isn't nulhomotopic by the last lemma. ■

8/18/2025

Today I want to start learning about general manifolds. So, I will switch over to following John Lee's book Introduction to Smooth Manifolds. Soon I might switch back either to Munkres or Guillemin.

Suppose M is a topological space. Then we say M is a topological manifold of dimension n (a.k.a a topological n-manifold) if:

- M is second countable and Hausdorff.
- M is locally Euclidean of dimension n, meaning that for any $p \in M$, there exists an open neighborhood $U \subseteq M$ of p and a homeorphism φ from U to an open set $\hat{U}$ of H^n or $\mathbb{R}^n$.

If $n = 0$, we say M is a topological 0-manifold if M is countable and equipped with the discrete topology.

While it's clear that each pair (U, φ) define local coordinates on M, one weird thing notation-wise is that those coordinates are commonly written as:

$\varphi(p) = (x^1(p), \dots, x^n(p))$ (as opposed to using subscripts).

This notation also extends to trivial Euclidean manifolds where (for example) John Lee denotes $x \in \mathbb{R}$ as $x = (x^1, \dots, x^n)$. I don't know why this is apparently common notation in this field of math. (Also, to be clear my differential geometry professor before I dropped the course also used that notation. So, I'm not making it up that the notation is common....)

One manifold I haven't seen before is the n-dimensional real projective space, denoted $\mathbb{RP}^n$. It is defined as the set of 1-dimensional linear subspaces of $\mathbb{R}^{n+1}$ equipped with the quotient topology determined by the map $\pi : \mathbb{R}^{n+1} \to \mathbb{RP}^n$ where $x \mapsto \operatorname{span}\{x\}$.

Note: Given $x \in \mathbb{R}^{n+1}$, denote $[x] \coloneqq \pi(x)$.

To prove that $\mathbb{RP}^n$ is in fact a manifold, for each $i \in \{1, \dots, n+1\}$ let $\tilde{U}_i = \{x \in \mathbb{R}^{n+1} : x^i \neq 0\}$. Then let $U = \pi(\tilde{U}_i)$. Now $\tilde{U}_i$ is a saturated set with respect to π (meaning $\pi^{-1}(\pi(U_i)) = U_i$) for each i. Hence $\pi(U_i)$ is open $\mathbb{RP}^n$ for each i.

Next, define a map $\varphi_i : U_i \to \mathbb{R}^n$ by:

$$\varphi([(x^1, \dots, x^{n+1})]) \coloneqq (\tfrac{x^1}{x^i}, \dots, \tfrac{x^{i-1}}{x^i}, \tfrac{x^{i+1}}{x^i}, \dots, \tfrac{x^{n+1}}{x^i}).$$

If we scale $x \in \mathbb{R}^{n+1}$ by a nonzero constant, $\varphi_i([x])$ doesn't change. It follows that φ_i is a well-defined map. Also, φ_i is continuous since it's clear that $\varphi_i \circ \pi$ is continuous and π is a quotient map. Finally, to show that φ_i has a continuous inverse, note that:

$$\varphi_i^{-1}(u^1, \dots, u^n) = [(u^1, \dots, u^{i-1}, 1, u^{i+1}, \dots, u^{n+1})].$$

Since φ_i^{-1} is the composition of a continuous function from $\mathbb{R}^n$ to $\mathbb{R}^{n+1}$ and π which is also continuous, we have that φ_i^{-1} is continuous. So φ_i is a homeomorphism.

Since the U_i cover $\mathbb{RP}^n$ we've thus shown that $\mathbb{RP}^n$ is locally Euclidean of dimension n. The proof that $\mathbb{RP}^n$ is second countable and Hausdorff is left as an exercise by Lee.

We know $\mathbb{RP}^n$ is second countable since it is the union of a finite number of open second countable sets.

To show that $\mathbb{RP}^n$ is Hausdorff, suppose that $[x_1], [x_2] \in \mathbb{RP}^n$ with $[x_1] \neq [x_2]$, and without loss of generality suppose x_1 and x_2 are unit vectors. Then x_1 and x_2 are not collinear. So, using the Hausdorffness of S^n, there is an open neighborhood $U_1 \subseteq S^n$ of x_1 such that $x_2, -x_2 \notin U_1$, and similarly there is an open neighborhood $U_2 \subseteq S^n$ of x_2 such that $x_1, -x_1 \notin U_2$.

Now define $V_i = \{cu : c \in \mathbb{R} - \{0\}, u \in U_i\}$ for each i. Then each V_i is easily checked to be a saturated open set with respect to π. Also, we have that $[x_2] \notin \pi(V_1)$ and $[x_1] \notin \pi(V_2)$. So $\pi(V_1)$ and $\pi(V_2)$ are disjoint open sets in $\mathbb{RP}^n$ separating $[x_1]$ and $[x_2]$.

One more note is that if we restrict π to just S^n, then we get a continuous surjective map from a compact set to $\mathbb{RP}^n$. This says that $\mathbb{RP}^n$ is compact. ■

Lemma 1.10: Every topological manifold M without boundary has a countable basis of precompact coordinate balls (i.e. sets that are homeomorphic to an open ball in $\mathbb{R}^n$) and coordinate half-balls (i.e. sets that are homeomorphic to $B_r(x) \cap H^n$ where $B_r(x)$ is a ball centered at a point in ∂H^n).

Proof:

Firstly, since every point in M has a coordinate patch (Lee uses the word "chart") about it, there exists an open covering $\{U_\alpha\}_{\alpha \in A}$ of M consisting of domains of coordinate patches $\varphi_\alpha : U_\alpha \to \hat{U}_\alpha \subseteq \mathbb{R}$.

Lemma: If X is a second countable space and $\{U_\alpha\}_{\alpha \in A}$ is an open cover of X, then there exists a countable subcover $\{U_{\alpha_n}\}_{n \in \mathbb{N}}$ of X.

Proof:

Let $\mathcal{B} = \{B_n\}_{n \in \mathbb{N}}$ be a countable basis for the topology of X. Then for each n we know there exists U_{α_n} with $B_n \subseteq U_{\alpha_n}$. Then in turn, since $\mathcal{B}$ is an open cover of X, we have that $\{U_{\alpha_n}\}_{n \in \mathbb{N}}$ also covers X.

It follows from that lemma there is a countable collection of coordinate charts $\varphi_{\alpha_n} : U_{\alpha_n} \to \hat{U}_{\alpha_n}$ covering M. Then, it's obvious how each φ_{α_n} defines a countable basis for U_{α_n} consisting of precompact coordinate balls and half balls, and if we take the union of all those bases for each n, we get a countable basis for all of M. ■

Corollary: Every topological manifold M is locally path connected (meaning there exists a basis of M consisting of path connected sets).

This is cause every coordinate ball is path connected on account of being homemorphic to a path connected set.

Now due to my patchy background, I'm only sorta familiar with how topological connectedness works. So, here is my attempt at reviewing / teaching some stuff to myself:

Here's what I already know / have proven in math 240B (Folland exercise 4.10):

- A topological space X is connected if there doesn't exist two nonempty open sets U and V in X such that $U \cap V = \emptyset$ and $U \cup V = X$.
- Equivalently, X is connected if the only two clopen sets are $\emptyset$ and X.

- If $\{E_\alpha\}_{\alpha \in A}$ is a collection of connected sets with nonempty intersection, then $E := \bigcup_{\alpha \in A} E_\alpha$ is connected.
- If $E \subseteq X$ is connected, then so is $\overline{E}$.
- For every $x \in X$ there is a maximal connected set $E \subseteq X$ containing x. This set is called a component of X. Additionally, we know that E is closed.
- A topological space X is path connected if for any $x, y \in X$, there is a continuous map $f : [0,1] \to X$ satisfying that $f(0) = x$ and $f(1) = y$.
- Any convex set in a topological vector space is path connected.

Here's some stuff I've used but not ever gotten around to proving before.

- If X is path connected, then X is connected.

  Proof:
  If X has only 1 element, then X is trivially both path-connected and connected.

  For the sake of contradiction, suppose X is path connected and that there exists nonempty open sets $U, V \subseteq X$ such that $U \cap V = \emptyset$ and $U \cup V = X$. Then pick $x \in U - V$ and $y \in V - U$. Since X is path connected, we know there is a path $f : [0,1] \to X$ from x to y.

  If we define $U' = f^{-1}(U)$ and $V' = f^{-1}(V)$, then the continuity of f guarentees that U' and V' are open subsets of $[0,1]$. Also, since U and V are disjoint and contain the entire range of f, we know that U' and V' are disjoint and their union is all of $[0,1]$. And since $0 \in U'$ and $1 \in V'$, neither sets are empty. All of this would points towards $[0,1]$ being a disconnected set.

  However, $[0,1]$ is connected.

  > Let U and V be two open sets partitioning $[0,1]$, and without loss of generality suppose $0 \in U$. Then set $\alpha := \sup\{s > 0 : [0,s) \subseteq U\}$. If $\alpha < 1$, then it'd be clear that $\alpha \notin U$. But it'd also be clear that α is not an interior point of $V = [0,1] - U$, which would contradict that V is open. So, we know that $\alpha = 1$. But now this requires that $[0,1] - U$ equals either $\{1\}$ or $\emptyset$, and only the latter is open in $[0,1]$. So there does not exist two open nonempty sets which form a partition of $[0,1]$.

  Hence, we have a contradiction and conclude that it is impossible for X to be path connected but not connected. ■

- If $f : X \to Y$ is continuous and X is connected, then $f(X)$ is connected. Similarly, if X is path connected, then $f(X)$ is path connected.

  Proof:
  If X is connected, let U and V be open sets in the subspace topology which partition $f(X)$. Then $f^{-1}(U)$ and $f^{-1}(V)$ are disjoint open subsets of X whose union is X. Since X is connected, it follows that either $f^{-1}(U)$ or $f^{-1}(V)$ is empty. And since $f(f^{-1}(U)) = U$ and $f(f^{-1}(V)) = V$ on account of the fact that both $U, V \subseteq f(X)$, we know either U or V is empty. So, $f(X)$ is connected.

If X is path connected, consider any $y_1, y_2 \in f(X)$. Then $\exists x_1, x_2 \in X$ with $f(x_i) = y_i$ for $i = 1, 2$, as well as a path $g : [0,1] \to X$ going from x_1 to x_2. It follows that $f \circ g$ is a path going from y_1 to y_2 in $f(X)$. ■

- If $\{E_\alpha\}_{\alpha \in A}$ is a collection of path connected sets with nonempty intersection, then $E := \bigcup_{\alpha \in A} E_\alpha$ is connected.

  Proof:
  Consider any $x, y \in E$, and let $z \in \bigcap_{\alpha \in A} E_\alpha$. Then there is path contained in one of the E_α going from x to z, and there is another path in another E_α going from z to y. Combining those paths gives us a path from x to y. ■

- E being path-connected does not necessarily mean that $\overline{E}$ is.

  Proof:
  Let $S = \{(t, \sin(1/t)) \in \mathbb{R}^2 : t > 0\}$. Then S is clearly path connected. Also, we can see that $\overline{S} = S \cup (\{0\} \times [-1,1])$. But, there is no continuous path going from any point in $\{0\} \times [-1,1]$ to any point in S. Hence $\overline{S}$ is not connected. ■

- For every $x \in X$ there is a maximal path connected set $E \subseteq X$ containing x. This set is called a path component of X.

  Proof:
  Let $\mathcal{E} = \{E_\alpha\}_{\alpha \in A}$ be the collection of every path connected set in X containing x. Note that $\mathcal{E}$ is not empty since $\{x\} \in \mathcal{E}$. Then $E = \bigcup_{\alpha \in A} E_\alpha$ is path connected. Also, clearly it is a maximal path connected set containing x. ■

- Let X be a topological space, and given $x \in X$ let E_x be the maximal connected component of X containing x. Then $\{E_x\}_{x \in X}$ forms a partition of X, meaning that if $E_x \cap E_y \neq \emptyset$, then $E_x = E_y$.

  Proof:
  Suppose $x, y \in X$ satisfy that $E_x \cap E_y \neq \emptyset$. Then $E_x \cup E_y$ is a connected subset containing both x and y. So, since E_x and E_y are maximal, we have that $E_x \cup E_y \subseteq E_x, E_y$. It follows that $E_x = E_x \cup E_y = E_y$. ■

- The previous bullet point also holds if we replace "maximal connected component" with "maximal path connected component".

- Two sets A and B are separated if $A \cap \overline{B} = \emptyset$ and $\overline{A} \cap B = \emptyset$. A topological space X is connected if and only if no two nonempty sets whose union is all of X are disconnected. (This is the definition in math 140a...)

  $(\Longrightarrow)$
  Let A and B be two nonempty separated sets satisfying that $A \cup B = X$. Then since $\overline{A} \cup B = X$ and $\overline{A} \cap B = \emptyset$, it's clear that $B = (\overline{A})^{\mathsf{C}}$. Similarly, it's clear that $A = (\overline{B})^{\mathsf{C}}$. So, both A and B are open. But now A and B are disjoint nonempty open sets partitioning X. This means that X is not connected.

$(\Longleftarrow)$
Let U and V be two disjoint nonempty open sets whose union is X. Then $V = U^C$ and $U = V^C$. This means U and V are closed and so $U = \overline{U}$, $V = \overline{V}$, and clearly both U and V are separated. ■

- If X is a topological space and $E \subseteq X$ is clopen, then E is a union of connected components.

  Proof:
  It suffices to show that if $x \in E$ and A_x is the connected component of X containing x, then $A_x \subseteq E$. Fortunately, since A_x is closed and E is clopen, we know that $A_x \cap E$ and $A_x - E$ are both closed subsets of X. In turn, we know that $A_x \cap E$ and $A_x - E$ are disjoint open sets in the relative topology of A_x whose union is all of A_x. Since A_x is connected, it follows that either $A_x \cap E = \emptyset$ or $A_x - E = \emptyset$. But the former case is not true since $x \in A_x \cap E$. So, we know that $A_x - E = \emptyset$. This shows that $A_x \subseteq E$. ■

Here's some nicher proofs:

- A space X is locally connected if X has a basis consisting of sets which are connected. Similarly, X is locally path connected if X has a basis consisting of sets which are path connected.

- If X is locally connected, then every connected component of X is open.

  Proof:
  Let E be a connected component of X. Then for all $y \in E$, there exists a connected open set U_y containing y. And since E is the maximal connected set containing y, we have that $U_y \subseteq E$. Hence, $E = \bigcup_{y \in E} U_y$ is open. ■

- By identical reasoning to the last bullet point, if X is locally path connected, then every path component of X is open.
- Suppose X is locally path connected. Then X is connected if and only if X is path connected.

  $(\Longrightarrow)$
  It suffices to show that X has only one path component. Luckily, if X had more than one, then since every path component is open, we'd be able to take the union of all but one in order to get two disjoint nonempty open sets whose union is all of X. But this contradicts that X is connected.

  $(\Longleftarrow)$
  We proved this direction several bullet points ago. ■

- If $\{X_\alpha\}_{\alpha \in A}$ is a family of connected topological spaces, then the product space $X = \prod_{\alpha \in A} X_\alpha$ is also connected.

  Proof:
  Let $<$ be a well-ordering of A, and for any $\beta \in A$ let $S_\beta = \{\alpha \in A : \alpha < \beta\}$ and $\overline{S_\beta} = S_\beta \cup \{\beta\}$. We shall proceed via transfinite induction.

Let $\langle x_\alpha \rangle_{\alpha\in A} \in \prod_{\alpha\in A} X_\alpha$. Then suppose that $\beta \in A$ satisfies that $\big(\prod_{\alpha\in\overline{S_\gamma}} X_\alpha\big) \times \big(\prod_{\alpha\in(\overline{S_\gamma})^{\mathsf{C}}} \{x_\alpha\}\big)$ is connected for every $\gamma \in S_\beta$.

We claim that $\big(\prod_{\alpha\in\overline{S_\beta}} X_\alpha\big) \times \big(\prod_{\alpha\in(\overline{S_\beta})^{\mathsf{C}}} \{x_\alpha\}\big)$ is connected.

To prove this, first note that if $\beta \in A$, then $X_\beta \times \prod_{\alpha\in A-\{\beta\}}\{x_\alpha\}$ is connected.

> Proof:
> Let U and V be disjoint open sets which partition $X_\beta \times \prod_{\alpha\in A-\{\beta\}}\{x_\alpha\}$, and let π_β be the projection of X onto X_β. Then $\pi_\beta(U)$ and $\pi_\beta(V)$ are disjoint open sets which partition X_β. It follows that one of those two sets is empty, and the only way that is possible is if either U or V is empty.

If $\overline{S_\beta} = \{\beta\}$, then this already proves our claim. Otherwise, notice that:

$$\Big(\prod_{\alpha\in\overline{S_\beta}} X_\alpha\Big) \times \Big(\prod_{\alpha\in(\overline{S_\beta})^{\mathsf{C}}} \{x_\alpha\}\Big) = \bigcup_{\gamma\in S_\beta} \Big(X_\beta \times \Big(\prod_{\alpha\in\overline{S_\gamma}} X_\alpha\Big) \times \Big(\prod_{\alpha\in(\overline{S_\gamma})^{\mathsf{C}}-\{\beta\}} \{x_\alpha\}\Big)\Big)$$

Now each of the sets in that union contain $\langle x_\alpha \rangle$. Also, we claim that each of them are connected. After all, note that:

$$X_\beta \times \Big(\prod_{\alpha\in\overline{S_\gamma}} X_\alpha\Big) \times \Big(\prod_{\alpha\in(\overline{S_\gamma})^{\mathsf{C}}-\{\beta\}} \{x_\alpha\}\Big) = \bigcup_{y\in X_\beta} \Big(\{y\} \times \Big(\prod_{\alpha\in\overline{S_\gamma}} X_\alpha\Big) \times \Big(\prod_{\alpha\in(\overline{S_\gamma})^{\mathsf{C}}-\{\beta\}} \{x_\alpha\}\Big)\Big)$$

We already know that $\{x_\beta\} \times \big(\prod_{\alpha\in\overline{S_\gamma}} X_\alpha\big) \times \big(\prod_{\alpha\in(\overline{S_\gamma})^{\mathsf{C}}-\{\beta\}} \{x_\alpha\}\big)$ is connected.

Also, for any $y \in X_\beta$, we can find a continuous map f from X to itself which sets the βth coordinate of a point to y and otherwise acts as the identity for all other coordinates. (We know this map is continuous because $\pi_\alpha \circ f$ is trivially continuous for all $\alpha \in A$...) This in turn shows that:

$$\{y\} \times \Big(\prod_{\alpha\in\overline{S_\gamma}} X_\alpha\Big) \times \Big(\prod_{\alpha\in(\overline{S_\gamma})^{\mathsf{C}}-\{\beta\}} \{x_\alpha\}\Big) \text{ is connected for all } y.$$

And since $X_\beta \times \prod_{\alpha\in A-\{\beta\}}\{x_\alpha\}$ is a connected set intersecting the set in the previous paragraph for all y and is a subset of the union:

$$\bigcup_{y\in X_\beta} \Big(\{y\} \times \Big(\prod_{\alpha\in\overline{S_\gamma}} X_\alpha\Big) \times \Big(\prod_{\alpha\in(\overline{S_\gamma})^{\mathsf{C}}-\{\beta\}} \{x_\alpha\}\Big)\Big),$$

we've thus shown that the entire union is connected. This finishes proving our claim at the top of this page.

By transfinite induction, we can now conclude that:

$$\Big(\prod_{\alpha\in\overline{S_\beta}} X_\alpha\Big) \times \Big(\prod_{\alpha\in(\overline{S_\beta})^{\mathsf{C}}} \{x_\alpha\}\Big) \text{ is connected for all } \beta \in A.$$

Finally, to finish our proof we can just note that:

$$X = \bigcup_{\beta\in A} \Big(\Big(\prod_{\alpha\in\overline{S_\beta}} X_\alpha\Big) \times \Big(\prod_{\alpha\in(\overline{S_\beta})^{\mathsf{C}}} \{x_\alpha\}\Big)\Big)$$

Also, every set in the above union is connected and contains $\langle x_\alpha \rangle_{\alpha \in A}$. Hence X is connected.

- If $\{X_\alpha\}_{\alpha \in A}$ is a family of path connected topological spaces, then the product space $X = \prod_{\alpha \in A} X_\alpha$ is also path connected.

  Proof:
  Let $\langle x_\alpha \rangle_{\alpha \in A}$ and $\langle y_\alpha \rangle_{\alpha \in A}$ be elements of X. Then we know for each $\alpha \in A$ that there is a path $f_\alpha : [0,1] \rightarrow X_\alpha$ such that $f_\alpha(0) = x_\alpha$ and $f_\alpha(1) = y_\alpha$. If we now define $f : [0,1] \rightarrow X$ by $f(t) = \langle f_\alpha(t) \rangle_{\alpha \in A}$, we will have that f is a continuous path from $\langle x_\alpha \rangle_{\alpha \in A}$ to $\langle y_\alpha \rangle_{\alpha \in A}$.
  (It is continuous because $\pi_\alpha \circ f = f_\alpha$ is continuous for all $\alpha \in A$.) ■

- Every quotient space of a connected space is connected. Also, every quotient space of a path connected space is path connected.

  Proof:
  Let X^* be a partition of X and let $f : X \rightarrow X^*$ be the function mapping every element to the set in X^* containing it. If we equip X^* with the quotient topology with resepct to f, then f will be continuous. Hence since X is connected, so will $f(X) = X^*$.

  Similar reasoning works when X is path connected. ■

- If X is a locally connected space, then every open set in X is locally connected. Similarly, if X is a locally path connected space, then so is every open set in X.

  This should be obvious.

- If $(X, \mathcal{T})$ is a topological space and $\mathcal{T}'$ is a coarser topology on $\mathcal{T}$, then $(X, \mathcal{T})$ being connected implies that $(X, \mathcal{T}')$ is connected. Similarly, $(X, \mathcal{T})$ being path connected implies that $(X, \mathcal{T}')$ is path connected. Another way of thinking about this is that adding sets to a topology only makes your space more disconnected.

  Proof:
  If $(X, \mathcal{T}')$ weren't connected, then the disjoint sets in $\mathcal{T}'$ partitioning X would also be an open partition in $(X, \mathcal{T})$.

  Next, if $f : [0,1] \rightarrow X$ is continuous with respect to $\mathcal{T}$, then we know it is also continuous with respect to $\mathcal{T}'$. This shows that a path with respect to $\mathcal{T}$ is also a path with respect to $\mathcal{T}'$. ■

8/22/2025

To start off today, I'm going to do an exercise from Folland.

Exercise 4.57: A collection $\mathcal{U}$ of open sets in X is called locally finite if each $x \in X$ has a neighborhood that intersects only finitely many members of $\mathcal{U}$. If $\mathcal{U}$ and $\mathcal{V}$ are open covers of X, $\mathcal{V}$ is a refinement of $\mathcal{U}$ if for each $V \in \mathcal{V}$ there exists $U \in \mathcal{U}$ with $V \subseteq U$. X is called paracompact if every open cover of X has a locally finite refinement.

Clearly any compact set is automatically paracompact since a finite subcover of an open cover will automatically be a locally finite refinement of that cover. However, beware that a refinement of a cover doesn't need to be a subset of the original cover, and it is possible for a cover to be locally finite without being finite.

The following exercise generalizes the theorems written on pages 81 and 82 of this journal.

(a) If X is a σ-compact LCH space, then X is paracompact. In fact, every open cover $\mathcal{U}$ has locally finite refinements $\{V_\alpha\}_{\alpha \in A}$ and $\{W_\alpha\}_{\alpha \in A}$ such that $\overline{V_\alpha}$ is compact and $\overline{W}_\alpha \subseteq V_\alpha$ for all $\alpha \in A$.

Let $(U_n)_{n\in\mathbb{N}}$ be an increasing sequence of precompact open sets such that $\overline{U_n} \subseteq U_{n+1}$ and $X = \bigcup_{n\in\mathbb{N}} U_n$. (Since X is a σ-compact LCH space, we proved in math 240b that such a sequence must exist...) Also, for ease of notation take $U_n = \emptyset$ whenever $n \leq 0$

Now, the collection $\{E \cap (U_{n+2} - \overline{U_{n-1}}) : E \in \mathcal{U}\}$ is an open cover of $\overline{U_{n+1}} - U_n$. And since $\overline{U_{n+1}} - U_n$ is compact, we can choose a finite subcover $\mathcal{V}_n$ from that collection. Doing this for all n and then setting $\mathcal{V} = \bigcup_{n\in\mathbb{N}} \mathcal{V}_n$, we have that $\mathcal{V}$ is a locally finite refinement of $\mathcal{U}$. After all, each x has a neighborhood $U_{n+2} - \overline{U_{n-1}}$ which only the finitely many open sets in $\mathcal{V}_n, \mathcal{V}_{n+1}, \mathcal{V}_{n+2}$, $\mathcal{V}_{n-1}$ and $\mathcal{V}_{n-2}$ can intercept. Also, each set in $\mathcal{V}$ is contained in some set of $\mathcal{U}$. And thirdly, we claim that if $V_\alpha \in \mathcal{V}$, then V_α is precompact. This is because V_α will be a closed subset of $\overline{U_{n+2}}$ for some n and the latter is compact.

Having constructed our first refinement $\mathcal{V} = \{V_\alpha\}_{\alpha\in A}$, we're now ready to construct our second. Fix n and note that each $x \in X$ has a compact neighborhood $N_x \subseteq V_\alpha$ for some $V_\alpha \in \mathcal{V}_n$. The N_x° form an open cover of $\overline{U_{n+1}} - U_n$. Hence, there is a finite collection $\{x_1, \ldots, x_m\}$ of points in $\overline{U_{n+1}} - U_n$ such that $\overline{U_{n+1}} - U_n \subseteq \bigcup_{j=1}^m N_{x_j}^\circ$. So, for each $\alpha \in A$ with $V_\alpha \in \mathcal{V}_n$ let W_α be the union of all the $N_{x_j}^\circ$ such that $N_{x_j} \subseteq V_\alpha$, and let $\mathcal{W}_n$ be the collection of W_α defined in this sentence. It's clear that $\mathcal{W}_n$ is also an open cover of $\overline{U_{n+1}} - U_n$, and that $\overline{W_\alpha} \subseteq V_\alpha$ for all $V_\alpha \in \mathcal{V}_n$. Doing this for all n, we then have that $\mathcal{W} = \bigcup_{n\in\mathbb{N}} \mathcal{W}_n$ is our second desired refinement.

(b) If X is a σ-compact LCH space, for any open cover $\mathcal{U}$ of X there is a partition of unity on X subordinate to $\mathcal{U}$ and consisting of compactly supported functions.

Let $\{V_n\}_{n\in\mathbb{N}}$ and $\{W_n\}_{n\in\mathbb{N}}$ be refinements of $\mathcal{U}$ constructed as in part (a). Note that our refinements in the last part were countable. So, there is no issue just taking the indexing set A to be $\mathbb{N}$.

For each n there exists by Urysohn's lemma a function $f_n \in C_c(X, [0,1])$ such that $f_n(x) = 1$ for all $x \in \overline{W_n}$ and $f_n(x) = 0$ for all $x \in V_n^{\mathsf{C}}$. Setting $f = \sum_{n=1}^\infty f_n$, we then have that f is continuous and finite everywhere in X on account of the fact that every $x \in X$ has a neighborhood where f is only a sum of finitely many continuous functions on that neighborhood. Also, since every $x \in X$ is contained in W_n, we know that $f(x) \geq 1$ for all x. Hence, if we define $g_n = f_n / f$ for each n, we have that g_n is well-defined and still in $C_c(X, [0,1])$.

We claim $(g_n)_{n\in\mathbb{N}}$ is a partition of unity subordinate to $\mathcal{U}$. After all, $\sum_{n\in\mathbb{N}} g_n = \frac{1}{f}\sum_{n\in\mathbb{N}} f_n = 1$. Also, each g_n satisfies that $\operatorname{supp}(g_n) \subseteq V_n$ and $V_n \subseteq U$ for some $U \in \mathcal{U}$. And finally, it is still the case that every $x \in X$ has a neighborhood on which only finitely many of the g_n are nonzero.

As a side note, we can extend this result to proving theorem 16.3 on page 82 of this journal by just noting that $\mathbb{R}^n$ is a σ-compact LCH space and that the Urysohn lemma on $\mathbb{R}^n$ specifies that we can choose each of our f_n to be in $C_c^\infty(X, [0,1])$.

Now I shall go over some more of John Lee's book.

Since manifolds are locally path connected, we know that a manifold is connected if and only if it is path connected. Also, it is clear that the path components of a manifold are identical to it's components. One more proposition is as follows:

Proposition 1.11.d: If M is a topological manifold, then M has countably many components, each of which are open subsets of M and a topological manifold by themselves.

Proof:
Since M is locally path connected, we know every single component is an open set. It follows that the components form an open cover $\mathcal{U}$ of M. And since M is second countable, this means that there is a countable subcover of $\mathcal{U}$. Yet, because all the sets of $\mathcal{U}$ are disjoint, the only way this is possible is if $\mathcal{U}$ was countable to begin with.

Also, it's clear that every component equipped with the subspace topology will still be second countable and Hausdorff. And to show that each component is locally Euclidean of dimension n, just restrict the domain and codomain of the coordinate patches covering that component. ■

Another consequence of every manifold having a countable basis of precompact coordinate balls and half balls is that manifolds are locally compact and σ-compact. And by the exercise I did earlier today, that means that every manifold is paracompact.

In fact, by slightly modifying our construction of $\mathcal{W}$ in the exercise I did before (namely by picking each N_x to be the closure of a precompact coordinate ball and then letting $\mathcal{W}_n$ consist of the $N_{x_j}^\circ$ without bothering to take any unions), we can construct a locally finite refinement consisting of coordinate balls and half balls for any open cover of M.

Here is a proposition about locally finite collections of sets.

Exercise 1.14: Suppose $\mathcal{X}$ is a locally finite collection of subsets of a topological space M.

(a) The collection $\{\overline{X} : X \in \mathcal{X}\}$ is also locally finite.

Proof:
Let $x \in M$ and let $U \subseteq M$ be an open set containing x that intersects only the elements of a finite subset $\mathcal{Y}$ of $\mathcal{X}$. Now we want to show that if $X \in \mathcal{X}$ satisfies $\overline{X} \cap U \neq \emptyset$, then $X \in \mathcal{Y}$. Luckily, if $\overline{X} \cap U \neq \emptyset$ but $X \cap U = \emptyset$ so that $X \notin \mathcal{Y}$, then it must be that any $x \in \overline{X} \cap U$ is an accumulation point of X. However, that is immediately contradicted by the fact that U is a neighborhood of x which does not intercept X anywhere. ■

(b) $\overline{\bigcup_{X \in \mathcal{X}} X} = \bigcup_{X \in \mathcal{X}} \overline{X}$.

Proof:
It is always true that $\bigcup_{X \in \mathcal{X}} \overline{X} \subseteq \overline{\bigcup_{X \in \mathcal{X}} X}$.

Meanwhile, to show the other inclusion, suppose x is an accumulation point of $\bigcup_{X \in \mathcal{X}} X$ but not in any individual $\overline{X}$. Then it must be the case that every neighborhood of x intercepts some set in $\mathcal{X}$, yet it must also be the case that for every $X \in \mathcal{X}$ there is a neighborhood of x with doesn't intercept X. The only way this is possible is if every neighborhood intercepts infinitely many $X \in \mathcal{X}$.

Let N be any neighborhood of x. Then we can construct an infinite sequence of distinct sets $(X_n)_{n \in \mathbb{N}}$ in $\mathcal{X}$ intercepting N as follows.

For the ease of notation set $U_0 = M$. Now at step n, choose any $X_n \in \mathcal{X}$ that intersects $N \cap \bigcap_{k=0}^{n-1} U_k$. Note that we can do this since a finite intersection of neighborhoods of x is still a neighborhood of x. But now we can choose some neighborhood U_n of x such that $X_n \cap U_n = \emptyset$. And now repeat this reasoning.

It's clear that all the chosen X_n intercept $\bigcap_{k=0}^{n-1} U_k \subseteq N$. Additionally, all the chosen X_n are distinct since $X_n \cap \bigcap_{k=0}^{N} U_k \subseteq X_n \cap U_n = \emptyset$ for all $N \geq n$.

But that contradicts that $\mathcal{X}$ is locally finite. ■

<u>Proposition 1.16:</u> The fundamental group of a topological manifold M is countable.

Proof:
Let $\mathcal{B}$ be a countable collection of coordinate balls and half balls covering M. For any $B, B' \in \mathcal{B}$, the intersection $B \cap B'$ has at most countably many components each of which is path connected.

Why? $B \cap B'$ is second countable and locally path connected on account of being an open subset of M. In turn, by the same reasoning as in proposition 1.11.d we know that $B \cap B'$ has only countably many components.

Let $\mathcal{X}$ be a countable set containing a point from each component of $B \cap B'$ for every $B, B' \in \mathcal{B}$ (including $B = B'$). Also, for each $B \in \mathcal{B}$ and $x, x' \in \mathcal{X}$ satisfying that $x, x' \in B$, let $h_{x,x'}^{B}$ be some path from x to x' in B. Since $\mathcal{X}$ intercepts every component of M, it suffices when calculating the fundamental group to take our base point p to be in $\mathcal{X}$. Then, we define a *special loop* to be a loop based at p that is a finite product of paths $h_{x,x'}^{B}$.

Now there are only countably many special loops. Therefore, in order to prove that $\pi_1(M,p)$ is countable, it suffices to show that if $f:[0,1]\to M$ is a loop based at p, then f is homotopic to some special loop.

Fortunately, the collection of the components of the sets $f^{-1}(B)$ with $B\in\mathcal{B}$ is an open cover of $[0,1]$. So, there exists $0=a_0<a_1<\cdots<a_k=1$ such that for each $i\geq 1$, $[a_{i-1},a_i]\subseteq f^{-1}(B)$ for some $B\in\mathcal{B}$. Now for each i, let f_i be the restriction of f to interval $[a_{i-1},a_i]$ and then reparametrized so that its domain is $[0,1]$, and also let $B_i\in\mathcal{B}$ be a coordinate ball containing the image of f_i. For each $1\leq i<k$, $f(a_i)\in B_i\cap B_{i+1}$. Also there is some $x_i\in\mathcal{X}$ that lies in the same component of $B_i\cap B_{i+1}$ as $f(a_i)$. So, let g_i be a path in $B_i\cap B_{i+1}$ from x_i to $f(a_i)$.

Note, we'll also write $x_0=x_k=p$ and $g_0=g_k$ are the constant paths based at p. Also, like in Munkres, we'll denote $\bar{g}_i$ to be the reverse path.

Now if we denote $\tilde{f}_i\coloneqq g_{i-1}*f_i*\bar{g}_i$, we have that:

$$\begin{aligned} f &\simeq_p f_1*\cdots*f_k \\ &\simeq_p g_0*f_1*\bar{g}_1*g_1*f_2*\bar{g}_2*g_2*\cdots*\bar{g}_{k-1}*g_{k-1}*f_k*\bar{g}_k \\ &\simeq_p \tilde{f}_1*\cdots*\tilde{f}_k. \end{aligned}$$

But now since each B_i is simply connected, we know that $\tilde{f}_i$ is path homotopic to $h^{B_i}_{x_{i-1},x_i}$. In turn, we have that f is path homotopic to a special loop. ■

A question I've had for a while is how smoothness can even be defined for manifolds which aren't subsets of $\mathbb{R}^n$. After all, differentiability as a concept relies on the topological structure of the reals. It turns out that the answer to this problem is to reframe some prior theorems as defining axioms which must be met.

Let M be a topological n-manifold. Two charts (U,φ) and (V,ψ) on M are said to be C^r compatible if either $U\cap V=\emptyset$ or the transition map $\psi\circ\varphi^{-1}$ from $\varphi(U\cap V)$ to $\psi(U\cap V)$ is a C^r diffeomorphism.

Note that the transition map is a function from an open open subset of either H^n or $\mathbb{R}^n$ to another open subset of either H^n or $\mathbb{R}^n$. So, it makes sense to talk about the transition map as being differentiable.

An atlas $\mathcal{A}$ for M is a collection of charts covering M. $\mathcal{A}$ is called a C^r atlas if any two charts in $\mathcal{A}$ are C^r compatible.

Very roughly speaking, if $1\leq s,r$, we want to define that a manifold M is C^r smooth if M has a C^r atlas $\mathcal{A}$. Additionally, we want to then say that a function $f:M\to\mathbb{R}$ is C^s differentiable if and only if $f\circ\varphi^{-1}$ is C^s differentiable for all charts $(U,\varphi)\in\mathcal{A}$.

A snag we need to work out though is that a space can have many C^r atlases. Also, some of those atlases may have charts which aren't C^r compatible with the charts in the other atlases. In a sense, that would mean they define different "smooth structures". At the same time, it could be the case that two atlases contain charts which are all C^r compatible with the charts in the other atlas. In that case, the two atlases could be said to define the same "smooth structure". We'll get around this issue as follows:

> We define a C^r atlas $\mathcal{A}$ on a topological manifold M to be maximal or complete if it is not contained in a larger C^r atlas. Or in other words, if a chart (U, φ) is C^r compatible with every chart in $\mathcal{A}$, then $(U, \varphi) \in \mathcal{A}$.
>
> If M is a topological manifold, then a C^r structure on M is a maximal C^r atlas.
>
> A C^r manifold is a pair $(M, \mathcal{A})$ where M is a topological manifold and $\mathcal{A}$ is a C^r structure on M.

Proposition 1.17: Let M be a topological manifold.

(a) Every C^r atlas $\mathcal{A}$ for M is contained in a unique maximal C^r atlas called the C^r structure determined by $\mathcal{A}$.

Proof:
Let $\mathcal{A}$ be a C^r atlas for M and define $\overline{\mathcal{A}}$ as the set of all charts which are C^r compatible with every chart in $\mathcal{A}$. If $\overline{\mathcal{A}}$ were a C^r atlas, it would be obvious that it is maximal. After all, if a chart was C^r compatible with every chart in $\overline{\mathcal{A}}$, then it would also be compatible with every chart in $\mathcal{A}$ on account of the fact that $\mathcal{A} \subseteq \overline{\mathcal{A}}$. But that would mean that that chart is also in $\overline{\mathcal{A}}$.

Hence, we proceed by trying to prove that $\overline{\mathcal{A}}$ is a C^r atlas on M. Or in other words, we want to show that for any $(U, \varphi), (V, \psi) \in \overline{\mathcal{A}}$, with $U \cap V \neq \emptyset$, the map $\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V)$ is smooth.

> Choose $x = \varphi(p) \in \varphi(U \cap V)$. Then there is some chart $(W, \theta) \in \mathcal{A}$ with $p \in W$. And since every chart in $\overline{\mathcal{A}}$ is C^r compatible with (W, θ), we know that $\theta \circ \varphi^{-1}$ and $\psi \circ \theta^{-1}$ are both C^r maps. It follows that $\psi \circ \varphi^{-1} = (\psi \circ \theta^{-1}) \circ (\theta \circ \varphi^{-1})$ is a C^r map on the neighborhood $\varphi(U \cap V \cap W)$ of x.
>
> This proves that $\phi \circ \varphi^{-1}$ is locally C^r. Hence it is C^r in general.

With that, we've prove that there exists a maximal C^r atlas $\overline{\mathcal{A}}$ containing $\mathcal{A}$. To finish off we show uniqueness. Suppose $\mathcal{B}$ is another maximal C^r atlas containg $\mathcal{A}$. Then every chart in $\mathcal{B}$ must be C^r compatible with every char in $\mathcal{A}$. But that then implies that $\mathcal{B} \subseteq \overline{\mathcal{A}}$. And since $\mathcal{B}$ is maximal, we have that $\mathcal{B} = \overline{\mathcal{A}}$. ■

(b) Two C^r atlases determine the same C^r structure iff their union is is a C^r atlas.

Proof:
Let $\mathcal{A}$ and $\mathcal{B}$ be C^r atlases and let $\overline{\mathcal{A}}$ and $\overline{\mathcal{B}}$ denote the smooth structures determined by $\mathcal{A}$ and $\mathcal{B}$ respectively.

$(\Longrightarrow)$
If $\overline{\mathcal{A}} = \overline{\mathcal{B}}$, then we know $\mathcal{B} \subseteq \overline{\mathcal{A}}$. So, every chart of $\mathcal{B}$ is C^r compatible with every chart of $\mathcal{A}$. It follows that any two charts in $\mathcal{A} \cup \mathcal{B}$ are smoothly compatible.

$(\Longleftarrow)$
If $\mathcal{A} \cup \mathcal{B}$ is a C^r atlas, then we know that $\mathcal{A} \cup \mathcal{B} \subseteq \overline{\mathcal{A}}$ and that $\mathcal{A} \cup \mathcal{B} \subseteq \overline{\mathcal{B}}$. It follows that both $\overline{\mathcal{A}}$ and $\overline{\mathcal{B}}$ are the unique smooth structure determined by $\mathcal{A} \cup \mathcal{B}$. So, $\mathcal{A} = \mathcal{B}$. ■

Note that in the prior reasoning, we took $r \in \mathbb{Z}_{>0} \cup \{\infty\}$. If $r = \infty$, we call a C^r manifold a smooth manifold.

Before doing the following exercise, I want to establish a useful result.

Proposition: Let M be a topological manifold. Also let $\mathcal{A}$ be a C^r atlas for M and $\overline{\mathcal{A}}$ be the C^r structure determined by $\mathcal{A}$.

(a) If (U, φ) is a chart in $\mathcal{A}$, then given any C^r diffeomorphism h acting on $\varphi(U)$, we have that the chart $(U, h \circ \varphi) \in \overline{\mathcal{A}}$.

Proof:
By the prior proposition it suffices to show that given another chart (V, ψ) in $\mathcal{A}$ such that $V \cap U \neq \emptyset$, $\psi \circ (h \circ \varphi)^{-1}$ and $(h \circ \varphi) \circ \psi^{-1}$ are C^r maps. Luckily, since h, h^{-1}, $\varphi \circ \psi^{-1}$, and $\psi \circ \varphi^{-1}$ are all C^r maps, this is obvious. ■

(b) If (U, φ) is a chart in $\mathcal{A}$, then given any open set $V \subseteq U$, $(V, \varphi|_V)$ is a chart in $\overline{\mathcal{A}}$.

Hopefully this is obvious.

The significance of the above lemma is that we can cut up and smoothly reparametrize a coordinate chart and we'll still get a chart which is in the C^r structure we've equipped our manifold with.

Problem 1-6: Let M be a nonempty topological manifold of dimension $n \geq 1$. If M has a C^r (or smooth) structure, then show that it has uncountably many distinct ones.

Proof:
We start by proving the following lemma...

Lemma: If $s > 0$, then $F_s(x) := \|x\|_2^{s-1} x$ defines a homeomorphism from B^n to itself and from $H^n \cap B^n$ to itself (where B^n is the open unit ball in $\mathbb{R}^n$). Also, F_s is a C^∞ diffeomorphism on $\mathbb{R}^n - \{0\}$, and F_s is also a C^k diffeomorphism (for any k) on $\mathbb{R}^n$ if and only if $s = 1$.

Proof:
It's clear that F_s is a continuous function on $\mathbb{R}^n$. Also if $x \neq 0$ satisfies that $\|x\|_2^{s-1} x = y$, then $x = \|x\|_2^{1-s} y$ and $\|y\|_2 = \|x\|_2^s$. In turn $\|y\|_2^{(1-s)/s} = \|x\|_2^{1-s}$ and we have thus derived the following inverse function for F_s:

$$F_s^{-1}(x) = \begin{cases} 0 & \text{if } x = 0 \\ \|x\|_2^{(1-s)/s} x & \text{if } x \neq 0 \end{cases}$$

We claim that F_s^{-1} is continuous. This is clearly true on $\mathbb{R}^n - \{0\}$. Meanwhile, to show that $\|x\|_2^{(1-s)/s} x \to 0$ as $x \to 0$, note that:

$$\left\| \|x\|_2^{(1-s)/s} x \right\|_2 = \|x\|_2^{1+\frac{1-s}{s}} = \|x\|_2^{1/s}.$$

The latter clearly goes to 0 as x goes to 0. So F_s^{-1} is continuous at 0. This proves that F_s is a homeomorphism from $\mathbb{R}^n$ to itself. To show that restricting F_s defines a homeomorphism on B^n or $H^n \cap B^n$ just requires noting that both F_s and F_s^{-1} map B^n into B^n and H^n into H^n. Also, since we have formulas for F_s and F_s^{-1}, we can now clearly see that F_s and F_s^{-1} are smooth on $\mathbb{R}^n - \{0\}$.

Finally, note that if $s = 1$, then $F_s(x) = x$ is clearly a smooth diffeomorphism. Meanwhile, if $s \neq 1$, then we either have that $s - 1 < 0$ or $(1-s)/s < 0$. In the former case, we know that F_s is not differentiable at 0. After all, given any $A \in L(\mathbb{R}^n, \mathbb{R}^n)$, we have that:

$$\frac{\|F_s(h) - F_s(0) - Ah\|_2}{\|h\|_2} = \frac{\|\|h\|^{s-1}h - Ah\|_2}{\|h\|_2} = \left\| (\|h\|^{s-1} I - A) \frac{h}{\|h\|_2} \right\|_2$$

It follows then that there is some sequence $(h_n)_{n \in \mathbb{N}}$ in $\mathbb{R}^n$ converging to 0 such that:

$$\left\| (\|h_n\|^{s-1} I - A) \frac{h_n}{\|h_n\|_2} \right\|_2 = \|(\|h_n\|^{s-1} I - A)\|_{\text{op}} \text{ for all } n.$$

I'll also note that by taking the negative of any neccesary h_n, we can force $(h_n)_{n \in \mathbb{N}}$ to be a sequence in H^n. I'm not actually sure if this is strictly necessary but who cares.

But now $\|(\|h_n\|^{s-1} I - A)\|_{\text{op}} \geq \left| \|h_n\|_{\text{op}}^{s-1} - \|A\|_{\text{op}} \right| \to \infty$ as $n \to \infty$ since $\|h_n\|^{s-1} \to \infty$ as $h_n \to 0$. This proves that the derivative of F_s at 0 doesn't exist when $s - 1 < 0$. Analogous reasoning shows that the derivative of F_s^{-1} at 0 doesn't exist when $(1-s)/s < 0$. So, F_s is not a diffeomorphism at 0 when $s \neq 1$.

Now let $\mathcal{A}$ be an atlas contained in our C^r structure for M. Then choose $p \in M$ and let (U, φ) be a chart in M containing p. By using the proposition I showed before doing this exercise, we can define another C^r atlas $\mathcal{A}'$ in the same structure on M such that (U, φ) is the only chart containing p. Specifically, define:

$$\mathcal{A}' := \{(V - \{p\}, \psi|_{V-\{p\}}) : (V, \psi) \in \mathcal{A}\} \cup \{(U, \varphi)\}.$$

Next, by restricting φ to a coordinate ball or half ball W centered at p and then reparametrizing, we can say that there is a chart (W, φ') in our structure on M satisfying that $p \in W$, $\varphi'(p) = 0$, and $\varphi'(W)$ equals either B^n or $H^n \cap B^n$. Hence, we define the C^r atlas:

$$\mathcal{A}'' := \{(V - \{p\}, \psi|_{V-\{p\}}) : (V, \psi) \in \mathcal{A}\} \cup \{(W, \varphi')\}$$

And now we're in a position to use our lemma. Given $s > 0$, let F_s be as in our lemma and define the atlas:

$$\mathcal{B}_s := \{(V - \{p\}, \psi|_{V-\{p\}}) : (V, \psi) \in \mathcal{A}\} \cup \{(W, F_s \circ \varphi')\}$$

Note that $\mathcal{B}_s$ is in fact an atlas for every s since F_s is a homeomorphism, meaning that $(W, F_s \circ \varphi')$ is a well-defined chart in M. Also, we can see that $\mathcal{B}_s$ is actually a C^r atlas. After all, we know from before that every pair of charts in $\mathcal{B}_s$ not including $(W, F_s \circ \varphi')$ are C^r compatible. Meanwhile, note that if $(V, \psi) \in \mathcal{A}$ satisfies that $(V - \{p\}) \cap W \neq \emptyset$, then F_s is a diffeomorphism on the set $\varphi'(V \cap W - \{p\})$ on account of $0 = \varphi'(p)$ not being in that set. It follows easily that $(F_s \circ \varphi') \circ \psi^{-1}$ defined on $\psi(V \cap W - \{p\})$ is a C^r diffeomorphism.

But now note that if $s \neq t$, then $\mathcal{B}_s$ and $\mathcal{B}_t$ do not generate the same C^r structure on M. After all, the charts $(W, F_s \circ \varphi')$ and $(W, F_t \circ \varphi')$ are not C^r compatible unless $s = t$. This is because when $x \neq 0$,

$$\begin{aligned}
((F_s \circ \varphi') \circ (F_t \circ \varphi')^{-1})(x) &= (F_s \circ \varphi' \circ (\varphi')^{-1} \circ F_t^{-1})(x) \\
&= (F_s \circ F_t^{-1})(x) \\
&= F_s(\|x\|_2^{(1-t)/t} x) \\
&= \|(\|x\|_2^{(1-t)/t})x\|_2^{s-1} \cdot \|x\|_2^{(1-t)/t} x \\
&= \|x\|_2^{\frac{(1-t)s}{t}} \|x\|_2^{s-1} x = \|x\|_2^{\frac{s}{t}-1} x = F_{s/t}(x).
\end{aligned}$$

Also, you can manually check that the transition map also equals $F_{s/t}(0)$ at $x = 0$. And since $F_{s/t}$ is a diffeomorphism of any class iff $s/t = 1$, we know that the two charts are C^r compatible if and only if $s = t$. ■

In a sense this exercise proves how important it is to keep in mind that we consider a manifold to be smooth with respect to a specific structure. That said, if we're not working with multiple different structures, then it's annoying to explicitely mention the structure over and over. So, we take the approach of calling a chart a smooth chart if it's in our structure.

To finish off today, I want to briefly address the boundary of a manifold.

Like before if M is a topological manifold, we say a point $p \in M$ is an interior point of M if there exists a chart (U, φ) covering p such that $\varphi(U)$ is open in $\mathbb{R}^n$. Meanwhile, if no such chart exists, we say p is a boundary point. Also, we denote the interior of M: $\operatorname{Int} M$, to be to be collection of interior points and the boundary of M: $\partial M := M - \operatorname{Int} M$.

It is easy to see that $\operatorname{Int} M$ is an open subset of M and a manifold by itself without a boundary. Based on that we can also easily see that ∂M is a closed subset of M.

Oh, I also forgot to mention: if $U \subseteq M$ is open and M is a C^r manifold, then we can view $U \subseteq M$ as a C^r submanifold of M. Specifically, it's clear that U is a second countable Hausdorff space in the relative topology. Also, given any smooth chart (V, φ) on M, we define a smooth chart $(U \cap V, \varphi|_{U \cap V})$ on U. This gives us a C^r structure on U.

8/24/2025

Let $r, k \in \mathbb{Z}_{>0} \cup \{\infty\}$ and always assume $k \leq r$.

Suppose M is a C^r manifold and $f : M \to \mathbb{R}^m$ is a function. We say f is a C^k function if for every $p \in M$ there exists a smooth chart (U, φ) such that $p \in U$ and $f \circ \varphi^{-1}$ is C^k on the open set $\varphi(U)$ of either H^n or $\mathbb{R}^n$. In this case we denote $f \in C^k(M, \mathbb{R}^m)$ (although when $m = 1$ we usually shorthand this as $f \in C^k(M)$).

Exercise 2.3: Let M be a C^r manifold and suppose $f : M \to \mathbb{R}^k$ is a C^k function where $k \leq r$. Show that $f \circ \psi^{-1} : \varphi(U) \to \mathbb{R}^m$ is C^k for every smooth chart (U, ψ) on M.

Proof:
Given any smooth chart (U, ψ), we can show that $f \circ \psi^{-1}$ is locally C^k as follows. Take any $x = \psi(p) \in \psi(U)$. Now we know there exists another smooth chart (V, φ) satisfying that $p \in V$ and that $f \circ \varphi^{-1}$ is C^k. Also, $\varphi \circ \psi^{-1}$ is a C^r map on $\psi(U \cap V)$. Thus, $f \circ \psi^{-1} = (f \circ \varphi^{-1}) \circ (\varphi \circ \psi^{-1})$ is a C^k map on the open neighborhood $\psi(U \cap V)$ of x in H^n or $\mathbb{R}^n$. ■

Side note: A recurring theme will be that you need M to be at to be at least C^k in order for C^k functions on M to be well-behaved. After all, if $r < k$ then the composite function above is no longer guarenteed to be C^k.

Corollary / Exercise 2.1: $C^k(M, \mathbb{R}^m)$ is a real vector space. If $m = 1$, then $C^k(M)$ is a a real commutative algebra.

We want to generalize the previous definition even more to cover maps from manifolds to other manifolds. To do this, it's worth noting that any open set U of either H^n or $\mathbb{R}^n$ can be thought of as a manifold. Furthermore, U is a C^r manifold for any r when equipped with the standard structure, i.e. the one determined by the atlas $\{(U, \mathrm{Id})\}$.

(Unless specified otherwise, always assume an open subset of $\mathbb{R}^n$ or H^n is equipped with the standard structure…)

Exercise 2.2: Let U be an open subset of H^n or $\mathbb{R}^n$. Then a function $f : U \to \mathbb{R}^m$ is C^k in the traditional real analysis definition iff it is C^k with respect to our new definition.

$(\Longrightarrow)$
Suppose (V, φ) is any smooth chart on U. Then $\varphi^{-1} = \mathrm{Id} \circ \varphi^{-1}$ is a C^r function. So, $f \circ \varphi^{-1}$ is C^k.

$(\Longleftarrow)$
It must be the case that $f \circ \mathrm{Id}^{-1} = f$ is C^k in the traditional real analysis sense since (U, Id) is a smooth chart. ■

Now note that when viewing $\mathbb{R}^m$ as being a C^r manifold, we can "symmetrize" our definition by noting that a function $f : M \to \mathbb{R}^m$ is C^k if and only if for all $p \in M$ there exists a smooth chart (U, φ) on M and another smooth chart (V, ψ) on $\mathbb{R}^m$ with $f(U) \subseteq V$ such that $\psi \circ f \circ \varphi^{-1}$ is C^k from $\varphi(U)$ into $\psi(V)$.

$(\Longrightarrow)$
Since f is C^k, let (U, φ) be a smooth chart such that $f \circ \varphi^{-1}$ is C^k. Then let (V, ψ) be any smooth chart on $\mathbb{R}^m$ with $f(U) \subseteq V$. Note that such a chart must exist since we know $(\mathbb{R}^m, \mathrm{Id})$ works. Now $\psi \circ \mathrm{Id}^{-1} = \psi$ is a C^r map from V. So, $\psi \circ f \circ \varphi^{-1}$ is a C^k from $\varphi(U)$ into $\psi(V)$.

$(\Longleftarrow)$
This direction is obvious when you just take (V, ψ) to be the chart $(\mathbb{R}^m, \mathrm{Id})$. ■

This points us to following generalization of differentiability on manifolds. Suppose M and N are both C^r manifolds and let $F : M \to N$ be any map. We say F is a C^k map if for every $p \in M$ there exists a smooth chart (U, φ) on M with $p \in U$ and another smooth chart (V, ψ) on N with $F(U) \subseteq V$ satisfying that the composite map $\psi \circ F \circ \varphi^{-1}$ is a C^k map from $\varphi(U)$ into $\psi(V)$.

Proposition 2.4: Every C^k map between two C^r manifolds M and N is continuous.

Proof:

Suppose $F : M \to N$ is C^k. Then given any $p \in M$, we can show that F is continuous on a neighborhood U of p. Specifically, let (U, φ) and (V, ψ) be smooth charts as in the prior definition. Then $\psi \circ F \circ \varphi^{-1}$ is continuous on the set $\varphi(U)$ on account of it being a differentiable function between two subsets of $\mathbb{R}^n$. Also, since both ψ and φ are homeomorphisms, we have that $F = \psi^{-1} \circ (\psi \circ F \circ \varphi^{-1}) \circ \varphi$ is a continuous map from $U \subseteq M$ into N.

Since each $p \in M$ has a neighborhood on which F is continuous, it follows that F is a continuous map from M to N. ■

Shit I just realized that I've never actually proven that continuity is local like that. So here's a quick lemma...

Lemma: Let $f : X \to Y$ be a map and suppose that every $x \in X$ has a neighborhood N_x such that $f|_{N_x}$ is continuous. Then f is continuous.

Proof:
For any $x \in X$, let N_x be a neighborhood satisfying that $f|_{N_x}$ is continuous. Then given any neighborhood V of $f(x)$ in Y, we know $U := f^{-1}(V) \cap N_x$ must be a neighborhood of x. satisfying that $f(U) \subseteq V$. So, f (not restricted to any subset) is continuous at x.

Since f is continuous at all $x \in X$, we know that f is continuous on X. ■

As a side note, when we were defining what it means for a map between manifolds, $F : M \to N$, to be differentiable, perhaps it felt overly restricting for us to force the chart (V, ψ) on N to satisfy that $F(U) \subseteq V$ in our definition. However, it turns out that without that requirement, it is no longer the case that F being differentiable implies that F is continuous.

Problem 2-1: Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$.

Now for every $x \in \mathbb{R}$ there are smooth coordinate charts (U, φ) containing x and (V, ψ) containing $f(x)$ such that $\psi \circ f \circ \varphi^{-1}$ is smooth as a map from $\varphi(U \cap f^{-1}(V))$ to $\psi(V)$. However f is clearly not continuous, nor smooth according to according to definition of smoothness of maps between two manifolds.

If $x \neq 0$, we can just pick $(U, \varphi) = (\mathbb{R} - \{0\}, \mathrm{Id})$ and $(V, \psi) = (\mathbb{R}, \mathrm{Id})$. Then it's clear that $f(x) \in V$ and $\psi \circ f \circ \varphi^{-1} = f$ is smooth as a map from $\varphi(U \cap f^{-1}(V)) = \mathbb{R} - \{0\}$ to $\mathbb{R}$.

Meanwhile, if $x = 0$, then pick $(U, \varphi) = (\mathbb{R}, \mathrm{Id})$ and $(V, \psi) = ((0, \infty), \mathrm{Id})$. Then it is still the case that $f(x) \in V$. Also, $\psi \circ f \circ \varphi^{-1} = f$ is just the constant function 1 on the set $\varphi(U \cap f^{-1}(V)) = [0, \infty)$. Thus since it can be extended to a differentiable function on an open set containing $[0, \infty)$, we can say that $\psi \circ f \circ \varphi^{-1}$ is also a smooth map from $\varphi(U \cap f^{-1}(V))$.

However, f is not differentiable or even continuous in the traditional analysis sense. Thus, f cannot be a smooth map from $\mathbb{R}$ as a manifold into $\mathbb{R}$ by exercise 2.2. Then in turn, we know from our prior efforts in generalizing differentiability that f is not smooth as a map into $\mathbb{R}$ as a manifold either. ■

Proposition 2.5: Suppose M and N are C^r manifolds, and $F : M \to N$ is a map. Then the following are equivalent.

(a) F is a C^k map.

(b) For every $p \in M$ there exist smooth charts (U, φ) containing p and (V, ψ) containing $F(p)$ such that $U \cap F^{-1}(V)$ is open in M and the map $\psi \circ F \circ \varphi^{-1}$ is C^k from $\varphi(U \cap F^{-1}(V))$ into $\psi(V)$.

(c) F is continuous and there exists atlases $\{(U_\alpha, \varphi_\alpha)\}_{\alpha \in A}$ of M and $\{(V_\beta, \psi_\beta)\}_{\beta \in B}$ of N consisting of smooth charts such that for each α and β:

$$\psi_\beta \circ F \circ \varphi_\alpha^{-1} \text{ is a } C^k \text{ map from } \varphi_\alpha(U_\alpha \cap F^{-1}(V_\beta)) \text{ into } \psi_\beta(V_\beta).$$

$(a \implies c)$
From the last proposition we know that F is continuous. Also, suppose (U, φ) and (V, ψ) are *any* smooth charts on M and N respectively such that $U \cap F^{-1}(V) \neq \emptyset$. Then we claim that $\psi \circ F \circ \varphi^{-1}$ is a C^k map from $\varphi(U \cap F^{-1}(V))$ into $\psi(V)$.

Proof:
Let $x = \varphi(p)$ be in $\varphi(U \cap F^{-1}(V))$. Then since F is C^k, we know there are smooth charts (W_m, θ_m) in M and (W_n, θ_n) in N such that $p \in W_m$, $F(W_m) \subseteq W_n$, and $\theta_n \circ F \circ \theta_m^{-1}$ is a C^k map from $\theta_m(W_m)$ into $\theta_n(W_n)$. Also, we have that $\psi \circ \theta_n^{-1}$ is a C^r map from $\theta_n(V \cap W_n)$ to $\psi(V \cap W_n)$. And similarly, we have that $\theta_m \circ \varphi^{-1}$ is a C^r map from $\varphi(U \cap W_m)$ to $\theta_m(U \cap W_m)$.

Now we get to composing the functions established above (and I'll do this slowly since my head is already spinning from all the symbols written above).

- $(\psi \circ \theta_n^{-1}) \circ (\theta_n \circ F \circ \theta_m^{-1})$ is a C^k map from $\theta_m(W_m \cap F^{-1}(V \cap W_n))$ into $\psi(V)$.
- Because $F^{-1}(W_n) \supseteq W_m$, we have that $W_m \cap F^{-1}(V \cap W_n) = W_m \cap F^{-1}(V)$.
- Thus $(\psi \circ \theta_n^{-1}) \circ (\theta_n \circ F \circ \theta_m^{-1}) \circ (\theta_m \circ \varphi^{-1})$ is a C^k map from $\varphi(U \cap W_m \cap F^{-1}(V))$ into $\psi(V)$.
- Also $\psi \circ F \circ \varphi^{-1} = (\psi \circ \theta_n^{-1}) \circ (\theta_n \circ F \circ \theta_m^{-1}) \circ (\theta_m \circ \varphi^{-1})$ and we know $x \in \varphi(U \cap W_m \cap F^{-1}(V))$.
- Since F is continuous, we know that $F^{-1}(V)$ is open. It follows that $\varphi(U \cap W_m \cap F^{-1}(V))$ is an open neighborhood of x in either H^n or $\mathbb{R}^n$.
- This shows that any $x \in \varphi(U \cap F^{-1}(V))$ has an open neighborhood in either H^n or $\mathbb{R}^n$ for which $\psi \circ F \circ \varphi^{-1}$ is a C^k map when restricted to that neighborhood. It follows that $\psi \circ F \circ \varphi^{-1}$ is C^k on $\varphi(U \cap F^{-1}(V))$.

Side note: I essentially just proved an analog of exercise 2.3 a few pages ago. So that I can cite it later, I'll write it out as follows...

<u>Proposition</u> If M and N are C^r manifolds and $F : M \to N$ is a C^k map, then $\psi \circ F \circ \varphi^{-1}$ is a C^k map on $\varphi(U \cap F^{-1}(V))$ for all smooth charts (U, φ) on M and (V, ψ) on N.

Based on the prior reasoning, it suffices to choose any covering of M and N of smooth charts and we are done showing (c).

$(c \implies b)$
Let $p \in M$ and let $(U_\alpha, \varphi_\alpha)$ be a smooth chart on M covering p. Next let (V_β, ψ_β) be a smooth chart on N covering $F(p)$. Since F is continuous, we know that $U_\alpha \cap F^{-1}(V_\beta)$ is open. Also, we know by assumption that $\psi_\beta \circ F \circ \varphi_\alpha^{-1}$ is C^k from $\varphi_\alpha(U_\alpha \cap F^{-1}(V_\beta))$ into $\psi_\beta(V_\beta)$. This proves (b).

$(b \implies a)$
Given any $p \in M$ let (U, φ) and (V, ψ) be as in the hypothesis of (b). Then since $U \cap F^{-1}(V)$ is open and contains p, we have that:

$(U', \varphi') := (U \cap F^{-1}(V), \varphi|_{U \cap F^{-1}(V)})$ is another chart on M containing p.

Importantly, $\psi \circ F \circ (\varphi')^{-1}$ will still be a C^k map from $\varphi'(U')$ into $\psi(V)$ since $\psi \circ F \circ \varphi^{-1}$ also is that. But, we also have $F(U') \subseteq V$. This proves (a). ■

Proposition 2.6: Let M and N be C^r manifolds, and let $F : M \to N$ be a map.

- If every point $p \in M$ has an open neighborhood U such that the restriction $F|_U$ is C^k, then F is C^k globaly.
- If F is C^k globaly, then it's restriction to every open subset $U \subseteq M$ is C^k.

Proof:
Hopefully it is obvious that the latter bullet point is just a corollary of the prosition I noted on the last page. Meanwhile, the first bullet point is proved just by noting that any smooth chart in U is also a smooth chart in M. ■

Corollary 2.8: (Gluing lemma) Let M and N be C^r manifolds and let $\{U_\alpha\}_{\alpha \in A}$ be an open cover for M. Suppose that for each $\alpha \in A$ we are given a C^k map $F_\alpha : U_\alpha \to N$ and suppose that $F_\alpha|_{U_\alpha \cap U_\beta} = F_\beta|_{U_\alpha \cap U_\beta}$ for all α and β. Then there exists a unique C^k map $F : M \to N$ such that $F|_{U_\alpha} = F_\alpha$ for all α.

Proof:
Define $F = \bigcup_{\alpha \in A} F_\alpha$. This a well defined function on M since the U_α cover all of M and for any $p \in M$ there is only one $q \in N$ such that $(p, q) \in F$. Also, it is clear that F is the unique map satisying that $F|_{U_\alpha} = F_\alpha$ for all $\alpha \in A$. And since all the F_α are C^k, we know that F is locally C^k. So by the last proposition, we know that F is a C^k map globally. ■

Proposition 2.10: Let M, N, and P be C^r manifolds.

(a) Every constant map $c : M \to M$ is C^r.

Proof:
Suppose $c(p) = q$ for all $p \in M$. Then let (V, ψ) be a smooth chart covering q. If (U, φ) is any smooth chart on M, we know that $c(U) = \{q\} \subseteq V$ and that $\psi \circ c \circ \varphi^{-1} = \psi(q)$ is a constant function on $\varphi(U)$. Therefore $\psi \circ c \circ \varphi^{-1}$ is C^r and we've proven that c is a C^r map. ■

(b) The identity map on M is C^r.

Proof:
Let (U, φ) be any chart on M. Then $\mathrm{Id}_M(U) \subseteq U$ and $\varphi \circ \mathrm{Id}_M \circ \varphi^{-1} = \mathrm{Id}_{\mathbb{R}^n}$ is C^r. This proves that Id_M is C^r.■.

(c) If $U \subseteq M$ is an open submanifold, then the inclusion map $U \hookrightarrow M$ is C^r.

Proof:
Just apply proposition 2.6 to the identity map on M. ■

(d) If $F : M \to N$ and $G : N \to P$ are C^k, then so is $G \circ F : M \to P$.

Proof:
Let $p \in M$. Then by definition there are smooth charts (V, θ) on N and (W, ψ) on P such that $F(p) \in V$, $G(V) \subseteq W$, and $\psi \circ G \circ \theta^{-1}$ is a C^k map on $\theta(V)$. Also, since F is continuous, we know that $F^{-1}(V)$ is an open set in M. Therefore, we can find a smooth chart (U, φ) such that $p \in U \subseteq F^{-1}(V)$. In turn, $(G \circ F)(U) \subseteq W$. And since F is C^k, we know that $\theta \circ F \circ \varphi^{-1}$ is a C^k map from $\varphi(U)$. Hence:

$$\psi \circ (G \circ F) \circ \varphi^{-1} = (\psi \circ G \circ \theta^{-1}) \circ (\theta \circ F \circ \varphi^{-1}) \text{ is a } C^k \text{ map from } \varphi(U)$$

This proves that $G \circ F$ is a C^k map. ■

8/25/2025

Today I'm going to jump back to Guillemin's Differential Forms. My reasoning for this is that I want a working fomulation of Stokes theorem before the end of the Summer, and at the rate I'm going through Lee's book, I'm not going to get that formulation from Lee. I also never quite finished the chapter on tensors before. I will continue to return to Lee off and on though.

The pullback operation on $\Lambda^k(V^*)$:
Let V be an n-dimensional spaces over a field F with characteristic 0, and let W be an m-dimensional vector space over F. Also let $A : V \to W$ be a linear map. Recall that for any $T \in \mathcal{L}^k(W)$ we defined $A^\dagger T(v_1, \ldots, v_k) = T(Av_1, \ldots, Av_k)$.

Lemma 1.8.1: If $T \in \mathcal{I}^k(W)$, then $A^\dagger T \in \mathcal{I}^k(V)$.

Proof:
Since T can be expressed as a linear combination of redundant k-tensors and $A^\dagger$ is a linear map from $\mathcal{L}^k(W)$ to $\mathcal{L}^k(V)$, it suffices to assume T is itself a redundant k-tensor. So let $T = \ell_1 \otimes \cdots \otimes \ell_k$ where each $\ell_j \in W^*$ and $\ell_i = \ell_{i+1}$ for some i. Then by proposition 1.3.18 (on page 103) we have that $A^\dagger T = (A^\dagger \ell_1) \otimes \cdots \otimes (A^\dagger \ell_k)$. It follows that $A^\dagger T$ is a redundant k-tensor. ■.

Let π_W and π_V be the projections of $\mathcal{L}^k(W)$ and $\mathcal{L}^k(V)$ onto $\Lambda^k(W^*)$ and $\Lambda^k(V^*)$ respectively.

If $\omega \in \Lambda^k(W^*)$ and $T \in \mathcal{L}^k(W)$ satisfies that $\pi_W(T) = \omega$, then we define:

$$A^\dagger \omega := \pi_V(A^\dagger T).$$

To see that this is well defined, suppose $\omega = \pi(T) = \pi(T')$. Then $T = T' + S$ for some $S \in \mathcal{I}^k(V)$. So, $A^\dagger T = A^\dagger T' + A^\dagger S$. And since $A^\dagger S \in \mathcal{I}^k(V)$ by the last lemma, we have that $\pi(A^\dagger T) = \pi(A^\dagger T')$.

Proposition 1.8.4: The map $A^\dagger : \Lambda^k(W^*) \to \Lambda^k(V^*)$ sending ω to $A^\dagger\omega$ is linear. Moreover:

- if $\omega_i \in \Lambda^{k_i}(W^*)$ for $i = 1, 2$, then $A^\dagger(\omega_1 \wedge \omega_2) = (A^\dagger\omega_1) \wedge (A^\dagger\omega_2)$;
- if U is a vector space and $B : U \to V$ is a linear map, then for $\omega \in \Lambda^k(W^*)$, $B^\dagger(A^\dagger\omega) = (AB)^\dagger\omega$.

Proof:
Firstly, let $\omega, \omega' \in \Lambda^k(W^*)$. Then if $T \in \pi_W^{-1}(\{\omega\})$ and $T' \in \pi_W^{-1}(\{\omega\})$, we have for any $\lambda, \lambda' \in F$ that $\pi_W(\lambda T + \lambda' T') = \lambda\omega + \lambda'\omega'$. And now, showing that $A^\dagger$ as a map from $\Lambda^k(W^*)$ is linear is as simple as noting that $\pi_V \circ A^\dagger$ is linear (where we view $A^\dagger$ as a map from $\mathcal{L}^k(W)$).

Next, let ω_1, ω_2 be as in the proposition statement. Then suppose T_1 and T_2 both satisfy that $\pi_W(T_1) = \omega_1$ and $\pi_W(T_2) = \omega_2$ (ignore my abuse of notation). Then:

$$\begin{aligned} A^\dagger(\omega_1 \wedge \omega_2) &= \pi_V(A^\dagger(T_1 \otimes T_2)) \\ &= \pi_V((A^\dagger T_1) \otimes (A^\dagger T_2)) \\ &= \pi_V((A^\dagger T_1)) \wedge \pi_V((A^\dagger T_2)) = (A^\dagger\omega_1) \wedge (A^\dagger\omega_2). \end{aligned}$$

Finally, let $\omega \in \Lambda^k(W^*)$ and choose $T \in \mathcal{L}^k(W)$ such that $\pi_W(T) = \omega$. Then if you squint you can see that:

$$B^\dagger(A^\dagger\omega) = \pi_V(B^\dagger(A^\dagger T)) = \pi_V((AB)^\dagger T) = (AB)^\dagger\omega.$$

One application of the pullback operation is that it gives us a way of defining determinants completely independently of any chosen basis. Specifically, let V be an n-dimensional vector space over a field F with characteristic 0, and suppose $A : V \to V$ is a linear map.

(If you want to be pedantic, everything that follows should work so long as F has a characteristic that makes it so that $n! \neq 0$ and $-1 \neq 1$...)

Since $\dim \Lambda^n(V^*) = \binom{n}{n} = 1$ and $A^\dagger : \Lambda^n(V^*) \to \Lambda^n(V^*)$ is linear, it must be that the map $A^\dagger$ is just multiplication by a constant. We denote this constant $\det(A)$ and call it the determinant of A. In other words, we define $\det(A) \in F$ to be the constant such that $A^\dagger\omega = \det(A)\omega$ for all $\omega \in \Lambda^n(V^*)$.

Proposition 1.8.7: If A and B are linear mappings of V into V, then $\det(AB) = \det(A)\det(B)$.

Proof:
Given any $\omega \in \Lambda^n(V^*)$:

$$\det(AB)\omega = (AB)^\dagger\omega = B^\dagger(A^\dagger\omega) = \det(B)A^\dagger\omega = \det(B)\det(A)\omega$$

It follows that $\det(A)\det(B) = \det(AB)$. ■

Proposition 1.8.8: Write $\mathrm{Id}_V : V \to V$ for the identity map. Then $\det(\mathrm{Id}_V) = 1$.

Proof:
Note that if $\omega \in \Lambda^k(V^*)$ for any k and $T \in \mathcal{L}^k(V)$ satisfies that $\pi_V(T) = \omega$, then $\mathrm{Id}^\dagger\omega = \pi_V(\mathrm{Id}^\dagger T) = \pi_V(T) = \omega$. This shows that for any k, $\mathrm{Id}^\dagger$ is the identity map on $\Lambda^k(V^*)$. Hence $\det(\mathrm{Id}) = 1$. ■

Corollary: If $A : V \to V$ is a surjective linear map, then $\det(A) \neq 0$.

Proof:
If A is surjective, then we know by the rank-nullity theorem that A has nullity 0. So, A is bijective and has an inverse A^{-1}. Then by our last two propositions, we know that $\det(A)\det(A^{-1}) = \det(\mathrm{Id}_V) = 1$. Thus, it cannot be that $\det(A) = 0$. ■

Proposition 1.8.9: If $A : V \to V$ is not surjective, then $\det(A) = 0$.

Proof:
Let W be the image of A. If A is not surjective, we know that $\dim W < n$ and thus $\Lambda^n(W^*) = \{0\}$. So, let $A = i_W B$ where i_W is the inclusion map $W \hookrightarrow V$ and let B be the map A with it's codomain restricted to W. Then by proposition 1.8.4 we have that $A^\dagger \omega = B^\dagger(i_W^\dagger \omega)$. But now note $i_W^\dagger \omega \in \Lambda^n(W^*) = \{0\}$. This means that $i_W^\dagger \omega = 0$ and we trivially have that $B^\dagger(i_W^\dagger) = 0$. It follows that $\det(A) = 0$. ■

We still need to show that this definition of the determinant agrees with the usual one. To do this we can first prove something slightly more general.

Suppose $A : V \to W$ is a linear map, and let $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ be bases of V and W respectively. Then let $e_1^*, \ldots, e_n^*$ and $f_1^*, \ldots, f_n^*$ be the corresponding dual bases. If $(a_{i,j})$ is the $n \times n$ matrix representing A with respect to our bases (i.e. $Ae_j = \sum_{i=1}^n a_{i,j} f_i$ for all j), then:

$A^\dagger f_j^* = \sum_{i=1}^n a_{j,i} e_i^*$ for all j (see claim 1.2.15 on page 102...).

In turn:

$$\begin{aligned} A^\dagger(f_1^* \wedge \cdots \wedge f_n^*) &= (A^\dagger f_1^*) \wedge \cdots \wedge (A^\dagger f_n^*) \\ &= (\sum_{i=1}^n a_{1,i} e_i^*) \wedge \cdots \wedge (\sum_{i=1}^n a_{n,i} e_i^*) = \sum_{1 \leq k_1, \ldots, k_n \leq n} a_{1,k_1} \cdots a_{n,k_n} (e_{k_1}^* \wedge \cdots \wedge e_{k_n}^*) \end{aligned}$$

Next, if the multi-index $I = (k_1, \ldots, k_n)$ is repeating, then $e_{k_1}^* \wedge \cdots \wedge e_{k_n}^* = 0$. (This is a consequence of the fact that the wedge product with respect to 1-tensors is anti-commutative). It follows that we can cancel out a bunch of terms in the sum and be left with:

$$A^\dagger(f_1^* \wedge \cdots \wedge f_n^*) = \sum_{\sigma \in S_n} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} (e_{\sigma(1)}^* \wedge \cdots \wedge e_{\sigma(n)}^*)$$

But now note that:

$$\begin{aligned} (e_{\sigma(1)}^* \wedge \cdots \wedge e_{\sigma(n)}^*) &= \pi_V(e_{\sigma(1)}^* \otimes \cdots \otimes e_{\sigma(n)}^*) \\ &= \pi_V((e_1^* \otimes \cdots \otimes e_n^*)^\sigma) = \mathrm{sgn}(\sigma)\pi_V(e_1^* \otimes \cdots \otimes e_n^*) = \mathrm{sgn}(\sigma) e_1^* \wedge \cdots \wedge e_n^*. \end{aligned}$$

So, we conclude that $A^\dagger(f_1^* \wedge \cdots \wedge f_n^*) = \sum\limits_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} (e_1^* \wedge \cdots \wedge e_n^*)$.

Letting $W = V$ and $f_i = e_i$ for all i, we in turn have shown that:

$$\det(A) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

8/26/2025

In physics yesterday we started talking about statistical mechanics, and that inspired me to try focusing on real analysis again. So, I will be returning to Folland for a bit. I think I'll start off where I left off at chapter 7.

Let X be an LCH space and $\mathcal{B}_X$ be the collection of Borel sets on X. At where we left off (on page 62), we had showed that the space $M(X)$ of complex Radon measures on $(X, \mathcal{B}_X)$ is a normed complex vector space when equipped with the norm $\mu \mapsto \|\mu\| = |\mu|(X)$. Also, we had shown that the map $\mu \mapsto I_\mu$ where $I_\mu(f) = \int f \mathrm{d}\mu$ is an isometric isomorphism from $M(X)$ to $C_0(X)^*$.

Exercise 7.8: Suppose that μ is a Radon measure on X. If $\phi \in L^1(\mu)$ and $\phi \geq 0$, then $\nu = \phi \mathrm{d}\mu$ is a Radon measure.

Since $\phi \in L^1(\mu)$, we know that $\nu(E) = \int_E \phi \mathrm{d}\mu$ is finite for all $E \in \mathcal{B}_X$. Thus ν is a finite measure. This trivially satisfies the requirement that $\nu(K)$ is finite for all compact K.

Now let $\varepsilon > 0$ and note that by corollary 3.6 (see my math 240a notes from Fall quarter), there exists $\delta > 0$ such that $\mu(A) < \delta$ implies that $|\nu(A)| = \nu(A) < \varepsilon$ for all $A \in \mathcal{B}_X$. This easily let's us show all the desired regularity properties of ν.

If $E \in \mathcal{B}_X$ then let $U \supseteq E$ be an open set such that $\mu(U - E) < \delta$. Then we know that $\nu(U - E) < \varepsilon$. Taking $\varepsilon \to 0$ shows that ν is outer regular on E.

If $U \subseteq X$ is open, then let $K \subseteq U$ be a compact set such that $\mu(U - K) < \delta$. Then we know that $\nu(U - K) < \varepsilon$. Taking $\varepsilon \to 0$ shows that ν is inner regular on U.

Corollary: If μ is a fixed positive Radon measure and $f \in L^1(\mu)$, then $\nu = f \mathrm{d}\mu$ is a complex Radon measure.

If we write $f = f_1 - f_2 + i(f_3 - f_4)$ and set $\nu_j = f_j \mathrm{d}\mu$ for all j, then it's clear from the last exercise that all the ν_j are finite (and thus complex) Radon measures. Also, $\nu = \nu_1 - \nu_2 + i(\nu_3 - \nu_4)$. So ν is also a complex radon measure.

In one sense the following is completely unnecessary. I will never need the full generality of what I'm about to prove (probably). On the other hand, Folland doesn't prove this and instead points his readers to the bibliography. So, for the first time in my life I went and procured a book from the bibliography of a math textbook. Also, this was especially a pain since when I finally got the textbook, it wouldn't convert to a pdf for some reason and my main E-book reader couldn't read it. So I eventually downloaded a DJVu reader that looks like it's from the fucking 2000s in order to finally read the book.

Considering the fact that there aren't just pdfs of this book floating around willy nilly on the internet, I should probably give a citation instead of just vaguely describing it. The book I will be briefly following along with is Hewitt and Stromberg's Real and Abstract Analysis. I've also made a bibiliography section at the end of the pdf where this book will have the honor of being the first citation.

A measure μ on $(X, \mathcal{M})$ is called decomposable if there is a family $\mathcal{F} \subseteq \mathcal{M}$ with the following properties:

(i) $\mu(F) < \infty$ for all $F \in \mathcal{F}$;

(ii) The members of $\mathcal{F}$ are disjoint and their union is X;

(iii) If $\mu(E) < \infty$, then $\mu(E) = \sum_{F \in \mathcal{F}} \mu(E \cap F)$;

(iv) If $E \subseteq X$ and $E \cap F \in \mathcal{M}$ for all $F \in \mathcal{F}$, then $E \in \mathcal{M}$.

Also, we call $\mathcal{F}$ a decomposition of $(X, \mathcal{M}, \mu)$.

Note that if μ is σ-finite, then we clearly have that μ is decomposable on $(X, \mathcal{M})$.

Lemma 19.26: Let $(X, \mathcal{M})$ be a measurable space and let μ and ν be arbitrary measures on $(X, \mathcal{M})$ such that $\mu(X) < \infty$ and $\nu \ll \mu$. Then there exists a set $E \in \mathcal{M}$ such that:

(i.) For all $A \in \mathcal{M}$ with $A \subseteq E$, either $\nu(A) = 0$ or $\nu(A) = \infty$. Also, if $\nu(A) = 0$, then so does $\mu(A) = 0$.

(ii.) ν is σ-finite on E^{C}.

Proof:
Consider the family:

$$\mathcal{B} := \{B \in \mathcal{M} : \forall C \in \mathcal{A},\ C \subseteq B \Longrightarrow \nu(C) = 0 \text{ or } \nu(C) = \infty\}.$$

Importantly, we know that $\mathcal{B} \neq \emptyset$ since $\emptyset \in \mathcal{B}$, and also that $\mu(B) \leq \mu(X) < \infty$ for all $B \in \mathcal{B}$. So, it is well defined to set $\alpha := \sup_{B \in \mathcal{B}} \mu(B)$. Next note that if $B, B' \in \mathcal{B}$, then $B \cup B' \in \mathcal{B}$.

Suppose $C \subseteq B \cup B'$ is measurable. Then $\nu(C) = \nu(C \cap B) + \nu(C \cap (B' - B))$. And since $C \cap B$ and $C \cap (B' - B)$ are both measurable subsets of sets in $\mathcal{B}$, we know that both have ν-measure zero or infinity. It follows that $\nu(C) \in \{0, \infty\}$.

We can thus construct a nondecreasing sequence of sets $(B_n)_{n \in \mathbb{N}}$ in $\mathcal{B}$ such that $\lim_{n \to \infty} \mu(B_n) = \alpha$. Setting $D = \bigcup_{n \in \mathbb{N}} B_n$, we have that $\mu(D) = \alpha$ and we can also show that $D \in \mathcal{B}$ using near identical reasoning as in the prior pink text.

Next we show that ν is semifinite on D^{C}. Suppose $F \in \mathcal{M}$ with $F \subseteq D$ and $\nu(F) = \infty$. Then for the sake of contradiction assume that $\nu(G)$ equals 0 or ∞ for every measurable subset $G \subseteq F$. It would follow that $F \cup D \in \mathcal{B}$. But then since $\nu(F) > 0$ and $\nu \ll \mu$, we'd know that $\mu(F) > 0$. Hence, $F \cup D$ would be a set in $\mathcal{B}$ with $\mu(F \cup D) > \alpha$ (and that inequality is strict since $\alpha < \infty$). Yet that contradicts how we defined α.

We furthermore show that ν is σ-finite on D^{C}. Let:

$$\mathcal{F} := \{F \in \mathcal{M} : F \subseteq D^{\mathsf{C}} \text{ and } \nu \text{ is } \sigma\text{-finite on } F\}.$$

Like before, there exists a nondecreasing sequence $(F_n)_{n \in \mathbb{N}}$ in $\mathcal{F}$ such that $\lim_{n \to \infty} \mu(F_n) = \sup_{F \in \mathcal{F}} \mu(F) =: \beta$. And if we set $F = \bigcup_{n \in \mathbb{N}} F_n$, then we clearly have that $F \in \mathcal{F}$ and $\mu(F) = \beta$. Our claim is that $\nu(F^{\mathsf{C}} \cap D^{\mathsf{C}}) = 0$.

Suppose not. Since ν is semifinite on D^{C}, we would have that there exists a measurable set $H \subseteq F^{\mathsf{C}} \cap D^{\mathsf{C}}$ with $0 < \nu(H) < \infty$. But since $\nu \ll \mu$, we'd have that $\mu(H) > 0$. It would then follow that $F \cup H \in \mathcal{F}$ and satisfies that $\mu(F \cup H) > \beta$. But this contradicts how we defined β.

It easily follows that ν is σ-finite on $F \cup (F^{\mathsf{C}} \cap D^{\mathsf{C}}) = D^{\mathsf{C}}$.

To finish off, let $\mathcal{G} := \{B \in \mathcal{M} : B \subseteq D \text{ and } \nu(B) = 0\}$. Now $\mathcal{G}$ is nonempty since $\emptyset \in \mathcal{G}$. So, it is well defined to set $\gamma := \sup_{B \in \mathcal{G}} \mu(B)$. Also, like before we have that if $B, B' \in \mathcal{G}$ then $B \cup B' \in \mathcal{G}$. So, we can once again take the union of a nondecreasing sequence of sets to get a set $G \subseteq D$ in $\mathcal{G}$ with $\mu(G) = \gamma$. And finally, set $E := D \cap G^{\mathsf{C}}$.

(i.) Since $E \subseteq D$, we know that any measurable $A \subseteq E$ satisfies that $\nu(A) \in \{0, \infty\}$. Also, if $\nu(A) = 0$ then we must have that $\mu(A) = 0$ since otherwise $G \cup A \in \mathcal{G}$ and $\mu(G \cup A) > \gamma$, which is a contradiction.

(ii.) $E^{\mathsf{C}} = D^{\mathsf{C}} \cup G$. And since ν is σ-finite on D^{C} and $\nu(G) = 0$, we have that ν is σ-finite on E^{C}. ■

Theorem 19.27: An Extension of the Lebesgue-Radon-Nikodym Theorem:

Let $(X, \mathcal{M}, \mu)$ be decomposable via the decomposition $\mathcal{F}$, and let ν be an arbitrary signed measure such that $\nu \ll \mu$. Then there exists an extended real $\mathcal{M}$-measurable function $f : X \to \overline{\mathbb{R}}$ such that $\nu(A) = \int_A f \mathrm{d}\mu$ for all $A \in \mathcal{M}$ for which μ is σ-finite. Also, we can take f to be finite on any F which on which ν is σ-finite. And if g is any extended real $\mathcal{M}$-measurable function such that $\nu(A) = \int_A g \mathrm{d}\mu$ when $\mu(A) < \infty$, then $f\chi_E = g\chi_E$ μ-a.e. for all $E \in \mathcal{M}$ on which μ is σ-finite.

Proof:

We shall first consider the simpler case where ν is a positive measure.

Now restricting μ and ν to subspaces of $(X, \mathcal{M})$ won't change that $\nu \ll \mu$. Consequently, by restricting μ and ν to any subspace $F \in \mathcal{F}$ and applying our prior lemma, we can conclude that there are sets $D_F, E_F \in \mathcal{M}$ such that $D_F \cap E_F = \emptyset$; $D_F \cup E_F = F$; ν is σ-finite on D_F; and all measurable $A \subseteq E_F$ satisfy that $\nu(A) = 0 \Longrightarrow \mu(A) = 0$ and $\nu(A) \in \{0, \infty\}$.

Note that if ν is σ-finite on all of F, we can just take $D_F = F$ and $E_F = \emptyset$.

And going one step further, if we restrict μ and ν to the subspace D_F of $(X, \mathcal{M})$, then we know by the typical Lebesgue-Radon-Nikodym theorem that there is a finite measurable nonnegative function $f_0^{(F)} : X \to [0, \infty)$ such that $\nu(A) = \int_A f_0^{(F)} \mathrm{d}\mu$ for all $A \in \mathcal{M}$ with $A \subseteq D_F$.

Now define f by pasting together the functions:

$$f|_F(x) := \begin{cases} f_0^{(F)}(x) & \text{if } x \in D_F \\ \infty & \text{if } x \in E_F \end{cases}$$

- To show that f is measurable, let $(a, \infty]$ be an open ray in $\overline{\mathbb{R}}$. Then:

$$f^{-1}((a, \infty]) \cap F = f|_F^{-1}((a, \infty]) = (f_0^{(F)})^{-1}((a, \infty]) \cup E_F.$$

  And since $f_0^{(F)}$ is measurable on the subspace D_F of $(X, \mathcal{M})$, we've thus proven that $f^{-1}((a, \infty]) \cap F \in \mathcal{M}$ for all F. It follows from the definition of a decomposition that $f^{-1}((a, \infty]) \in \mathcal{M}$. And since the open rays form a basis for $\mathcal{B}_{\overline{\mathbb{R}}}$, we have proven that f is measurable.

- To show that $\nu(A) = \int_A f \mathrm{d}\mu$ when μ is σ-finite on A, first suppose that $\mu(A) < \infty$. Then we know from axiom (iii) of the definition of a decomoposition that there exists a countable subset $\mathcal{F}_0 \subseteq \mathcal{F}$ such that $\mu(A) = \sum_{F \in \mathcal{F}_0} \mu(A \cap F)$. One consequence of this is that:
$$\int_A f \mathrm{d}\mu = \sum_{F \in \mathcal{F}_0} \int_{A \cap F} f \mathrm{d}\mu.$$
Another consequence is that because the $F \in \mathcal{F}$ partition X, we know $\mu(A \cap (\bigcup_{F \in \mathcal{F}_0^{\mathsf{c}}} F)) = 0$. Then in turn, since $\nu \ll \mu$, we also know that $\nu(A \cap (\bigcup_{F \in \mathcal{F}_0^{\mathsf{c}}} F)) = 0$. So, we must have that $\nu(A) = \sum_{F \in \mathcal{F}_0} \nu(A \cap F)$.
And now we claim for each F that $\nu(A \cap F) = \int_{A \cap F} f \mathrm{d}\mu$.
    Clearly $\nu(A \cap F) = \nu(A \cap D_F) + \nu(A \cap E_F) = \int_{A \cap D_F} f \mathrm{d}\mu + \nu(A \cap E_F)$.

    Also, because $A \cap E_f \subseteq E_F$, we know that $\nu(A \cap E_F) \in \{0, \infty\}$ with $\mu(A \cap E_F) = 0$ if and only if $\nu(A \cap E_f) = 0$. So if $\mu(A \cap E_F) = 0$, then $\int_{A \cap E_F} f \mathrm{d}\mu = 0 = \nu(A \cap E_F)$. Meanwhile, if $\mu(A \cap E_F) > 0$, then $\int_{A \cap E_F} f \mathrm{d}\mu = \int_{A \cap E_F} (\infty) \mathrm{d}\mu = \infty = \nu(A \cap E_F)$. Either way, we have that:
$$\nu(A \cap F) = \int_{A \cap D_F} f \mathrm{d}\mu + \int_{A \cap E_F} f \mathrm{d}\mu = \int_A f \mathrm{d}\mu$$

    So, we've shown that $\nu(A) = \sum_{F \in \mathcal{F}_0} \int_{A \cap F} f \mathrm{d}\mu = \int_A f \mathrm{d}\mu$ when $\mu(A) < \infty$. The case where $\mu(A) = \infty$ and μ is σ-finite on A then easily follows.

- Finally, suppose g is as in the theorem statement. Then for every $F \in \mathcal{F}$ and every measurable set $A \subseteq D_F$, we have that $\nu(A) = \int_A f \mathrm{d}\mu = \int_A g \mathrm{d}\mu$. The only way this is possible is if $f = g$ μ-a.e. on D_F.

    Meanwhile, we know for any $F \in \mathcal{F}$ that $A := E_F \cap g^{-1}([0, \infty)) \in \mathcal{M}$ because g is measurable. Now suppose $\mu(A) > 0$. Then letting $A_n := \{x \in A : g(x) < n\}$ for each n, we have that the A_n form an increasing sequence of sets whose union is A. So, we know that $\nu(A) = \lim_{n \to \infty} \nu(A_n)$. Also, since $\nu(A) > 0$ we know that $\nu(A_n) > 0$ for some n. Then in turn $\nu(A_n) = \infty$ since $A_n \subseteq E_F$. But this contradicts the fact that $\nu(A_n) = \int g \mathrm{d}\mu \leq n\mu(A_n) < \infty$. So, we conclude that $\mu(A) = 0$.

    As a result, we've shown for any $F \in \mathcal{F}$ that $f = g$ μ-a.e. on $D_F \cup E_F = F$.

    The fact that $g = f$ μ-a.e. on any measurable set such that $\mu(E) < \infty$ is then a simple consequence of the fact that there is a countable subset $\mathcal{F}_0 \subseteq \mathcal{F}$ as well a μ-null set $N \subseteq X$ such that $E = N \cup (\bigcup_{F \in \mathcal{F}_0} (F \cap E))$. And if μ is σ-finite on E, then we can show that $g = f$ μ-a.e. on E by considering E as a countable union of sets on which we've already showed $g = f$ μ-a.e.

Now we come back to the case where ν is signed. Let ν^+ and ν^- be the positive and negative variations of ν, and also let P and N be measurable subsets of X such that $\nu^+(N) = 0$ and $\nu^-(P) = 0$. Then by our prior reasoning we know there exists measurable functions f^+ and f^- such that $\nu^+(A) = \int_A f^+ \mathrm{d}\mu$ and $\nu^-(A) = \int_A f^- \mathrm{d}\mu$ for all A on which μ is σ-finite. By setting $f = f^+ - f^-$, we thus get a measurable function such that $\nu(A) = \int_A f \mathrm{d}\mu$ for all A on which μ is σ-finite.

Now suppose g is as in the theorem statement, and let g^+ and g^- be it's positive and negative parts. Given any $F \in \mathfrak{F}$ we must have that $\int_A g d\mu = \nu(A) = \nu^+(A) \geq 0$ for all measurable $A \subseteq F \cap P$. It follows that $g \geq 0$ μ-a.e. on $F \cap P$, and in turn:

$\int_A g^+ d\mu = \nu^+(A)$ when $A \subseteq F \cap P$.

By similar reasoning, we can show that $g \leq 0$ μ-a.e. on $F \cap N$. This is important because it shows that $g^+ = 0$ μ-a.e. on $F \cap N$. So $\int_A g^+ d\mu = 0 = \nu^+(A)$ for all $A \subseteq F \cap N$. And hence, we can conclude that $\int_A g^+ = \nu^+(A)$ for all $A \subseteq F$. It easily follows that $\int_A g^+ = \nu^+(A)$ for all A satisfying that $\mu(A) < \infty$. But then by the prior reasoning we did, we know that $g^+ = f^+$ μ-a.e. on any set E which μ is σ-finite on.

Analogous reasoning can be used to show that $g^- = f^-$ μ-a.e. on any set E which μ is σ-finite on. ■

Corollary: Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space and suppose ν is any arbitrary signed measure such that $\nu \ll \mu$. Then there exists a measurable function $f : X \to \overline{\mathbb{R}}$ such that $\nu = f d\mu$. Also, if $g : X \to \overline{\mathbb{R}}$ is another function satisfying that $\nu = g d\mu$, then $f = g$ μ-a.e.

We can of course also write a version of the prior theorem and corollary for when ν is a complex measure. Specifically, just apply the prior theorem to the real and imaginary parts of ν separately. That said, the corollary stops being interesting if ν is complex.

Now, it's unfortunate that this extension of the Lebesgue-Radon-Nikodym theorem loses the ability to decompose ν into a continuous part and a mutually singular part. That said, an extremely convenient fact which makes the last theorem feel more worthwhile is that it turns out that every positive Radon measure is decomposable (with an asterisk attached).

To start off, we need an exercise from Folland.

Note that this exercise references a bunch of exercises which I already did in my Math 240a notes. Look there if you want to see stuff like the definition of the saturation of a measure. I even just now went back and clean up a bunch of problems with my answer for exercise 1.16(e)!! (haha why was I so stupid I swear I want to put an evoker to my head...)

Exercise 1.22: Let $(X, \mathcal{M}, \mu)$ be a measure space and let μ^* be the outer measure induced by treating μ as a premeasure on $\mathcal{M}$. Then let $\mathcal{M}^*$ be the σ-algebra of μ^*-measurable sets and $\overline{\mu} := \mu^*|_{\mathcal{M}^*}$.

(a) If μ is σ-finite, then $(X, \mathcal{M}^*, \overline{\mu})$ is the the completion of $(X, \mathcal{M}, \mu)$.

If $E \in \mathcal{M}^*$, then since μ is σ-finite we know by exercise 1.18 that there exists a set A in $\mathcal{M}_{\sigma\delta} \subseteq \mathcal{M}$ such that $E \subseteq A$ and $\overline{\mu}(A - E) = 0$. Hence $E = A \cup N$ where $A \in \mathcal{M}$ and $\overline{\mu}(N) = 0$. And since $\overline{\mu}$ is the restriction of the outer measure induced by μ, we know there is a null set $N' \in \mathcal{M}$ with $N \subseteq N'$. This proves that $\mathcal{M}^*$ is a subset of the completion of $\mathcal{M}$ (which I will hereafter call $\mathcal{M}^\wedge$).

Meanwhile, by Carathéodory's theorem we know that $(X, \mathcal{M}^*, \overline{\mu})$ is a complete measure space. So, suppose $E = F \cup N$ where $F \in \mathcal{M}$ and $N \subseteq N'$ with $\mu(N') = 0$. Then since $N' \in \mathcal{M}^*$ and $\overline{\mu}(N') = \mu(N') = 0$, we know $N \in \mathcal{M}^*$. In turn this means that $E = F \cup N \in \mathcal{M}^*$ and we've showed that $\mathcal{M}^\wedge \subseteq \mathcal{M}^*$.

Combining the last two paragraphs, we have that $\mathcal{M}^* = \mathcal{M}^\wedge$. And since there is only one measure that extends μ to $\mathcal{M}^\wedge$, we automatically have that $\overline{\mu}$ is the completion of μ.

(b) In general, $\overline{\mu}$ is the saturation of the completion of μ.

For the sake of notation, I will write that $\mathcal{M}^\wedge$ and $\mu^\wedge$ are the completions of $\mathcal{M}$ and μ respectively; that $\widetilde{\mathcal{M}}^\wedge$ is the collection of $\mu^\wedge$-locally measurable sets; and that $\widetilde{\mu}^\wedge$ is the saturation of of $\mu^\wedge$.

Our first goal is to show that $\widetilde{\mathcal{M}}^\wedge = \mathcal{M}^*$. Equivalently, this means we need to show that a set $E \subseteq X$ is μ^*-measurable if and only if it is locally $\mu^\wedge$-measurable.

$(\Longrightarrow)$
Suppose $E \subseteq X$ is μ^*-measurable and let $A \subseteq X$ be any set in $\mathcal{M}^\wedge$ with $\mu^\wedge(A) < \infty$. In part (a) we were able to show without using any σ-finiteness that $\mathcal{M}^\wedge \subseteq \mathcal{M}^*$. Thus, A and $E \cap A$ are μ^*-measurable. Also note that by definition of the completion of a measure, we can pick a set $F \in \mathcal{M}$ with $A \subseteq F$ and $\mu^\wedge(A) = \mu(F)$.

Now we claim by a similar argument as in exercise 1.18 that there exists a set $B \in \mathcal{M}$ such that $E \cap A \subseteq B$ and $\mu^*(B - (A \cap E)) = 0$.

Using exercise 1.18(a), for each $j \in \mathbb{N}$ pick $B_j \in \mathcal{M}$ such that $(A \cap E) \subseteq B_j$ and $\mu^*(B_j) \leq \mu^*(E \cap A) + \frac{1}{j}$. Then since $E \cap A$ is μ^*-measurable and $B \cap (E \cap A) = E \cap A$, we have that:

$$\mu^*(E \cap A) + \mu^*(B_j - (E \cap A)) = \mu^*(B_j) \leq \mu^*(E \cap A) + \tfrac{1}{j}$$

Next note that $\mu^*(E \cap A) \leq \mu^*(A) \leq \mu^*(F) = \mu(F) = \mu^\wedge(A) < \infty$. Thus, we can subtract $\mu^*(E \cap A)$ from both sides of our inequality above to get that:

$$\mu^*(B_j - (E \cap A)) < \tfrac{1}{j}$$

And now, if we define $B = \bigcap_{j \in \mathbb{N}} B_j$, it's clear that $B \in \mathcal{M}$ and that $\mu^*(B - (E \cap A)) \leq \mu^*(B_j - (E \cap A)) < \frac{1}{j}$ for all $j \in \mathbb{N}$. So, $\mu^*(B - (E \cap A)) = 0$.

Now since $\mu^*(B - (A \cap E)) = 0$, we know that there exists a set $C \in \mathcal{M}$ with $\mu(C) = 0$ and $B - (A \cap E) \subseteq C$.

Specifically, for each $n \in \mathbb{N}$ there exists a countable covering $\{C_j^{(n)}\}_{j \in \mathbb{N}}$ of $B - (A \cap E)$ such that $\mu(\bigcup_{j \in \mathbb{N}} C_j^{(n)}) \leq \sum_{j=0}^{\infty} \mu(C_j^{(n)}) < 1/n$ In turn, $C := \bigcap_{n \in \mathbb{N}} (\bigcup_{j \in \mathbb{N}} C_j^{(n)})$ is a set in $\mathcal{M}$ with $B - (A \cap E) \subseteq C$ and $\mu(C) = 0$.

Finally, we have that $N := B - (A \cap E) \in \mathcal{M}^\wedge$ because $N \subseteq C$ and $(X, \mathcal{M}^\wedge, \mu^\wedge)$ is complete. And since $B \in \mathcal{M} \subseteq \mathcal{M}^\wedge$, we have that $(A \cap E) = B - N \in \mathcal{M}^\wedge$. This proves that E is locally $\mu^\wedge$-measurable.

$(\Longleftarrow)$
Suppose $E \subseteq X$ is locally $\mu^\wedge$-measurable and then choose any $F \subseteq X$. If $\mu^*(F) = \infty$, then we trivially have that $\mu^*(F \cap E) + \mu^*(F - E) \leq \mu^*(F)$. So, it suffices to assume that $\mu^*(F) < \infty$. But then by exercise 1.18(a) there exists for each $j \in \mathbb{N}$ a set $A_j \in \mathcal{M}$ such that $F \subseteq A_j$ and $\mu(A_j) < \mu^*(F) + 1/j$. And, by taking the intersection of all the A_j we get a set $A \in \mathcal{M}$ such that $A \subseteq F$ and $\mu(A) = \mu^*(F) < \infty$.

Since E is locally $\mu^\wedge$-measurable, it follows that $A \cap E \in \mathcal{M}^\wedge$. Then since $\mathcal{M}^\wedge \subseteq \mathcal{M}^*$, we know that $\mu^*(F) = \mu^*(F \cap (A \cap E)) + \mu^*(F - (A \cap E))$. And this shows that E is μ^*-measurable since $F \cap (A \cap E) = F \cap E$ and $F - (A \cap E) = F - E$ on account of the fact that $F \subseteq A$.

With that, we've now shown that $\widetilde{\mathcal{M}}^\wedge = \mathcal{M}^*$. Also, since there is only one extension of μ to $\mathcal{M}^\wedge$ and both $\overline{\mu}$ and $\widetilde{\mu}^\wedge$ extend μ to $\mathcal{M}^\wedge$, we must have that $\overline{\mu}(E) = \widetilde{\mu}^\wedge(E)$ whenever $E \in \mathcal{M}^\wedge \subseteq \mathcal{M}^*$ Also, by definition of the saturation of a measure, we have that if $E \in \mathcal{M}^* - \mathcal{M}^\wedge$, then $\widetilde{\mu}^\wedge(E) = \infty$. So, all we need to do left is show that if E is locally $\mu^\wedge$-measurable but not in $\mathcal{M}^\wedge$, then $\overline{\mu}(E) = \mu^*(E) = \infty$.

Suppose E is μ^*-measurable and $\overline{\mu}(E) = \mu^*(E) < \infty$. When we were proving the backwards implication before, we showed that there is a set $A \in \mathcal{M}$ with $A \subseteq E$ and $\mu^*(E) = \mu(A) < \infty$. Next, when we were proving the forwards implication, we showed that there is a set $B \in \mathcal{M}$ with $A \cap E = E \subseteq B$ and $\mu^*(B - (A \cap E)) = \mu^*(B - E) = 0$. Then afterwords, we showed that there exists $C \in \mathcal{M}$ with $\mu(C) = 0$ and $B - E \subseteq C$. Finally, we have that $B - E \in \mathcal{M}^\wedge$ and in turn that $E = B - (B - E) \in \mathcal{M}^\wedge$. ■

As a side note, this shows that you can't indefinitely extend a measure space to larger and larger σ-algebras just by applying the theorem on page 24 of my latex math 240a notes over and over again. After all, the saturation of a complete measure is still complete (by exercise 1.16 in my latex math 240a notes). So, if you complete it again you just get back the same set. Also, it's easy to see that saturating a measure does not add any new locally measurable sets since all the added measurables sets have infinite measure. So, taking a saturation twice is the same as taking it once.

Now let X be an LCH space and let μ be a Radon measure on $(X, \mathcal{B}_X)$. Then set:

$$\mu^*(E) \coloneqq \inf\{\mu(U) : U \supseteq E \text{ with } U \text{ open}\} \text{ for all } E \in \mathcal{P}(X).$$

Recall from the proof of the Riesz Representation thorem in my math 240c notes that μ^* is a well-defined outer measure for which every Borel set is μ^*-measurable and $\mu^*|_{\mathcal{B}_X} = \mu$. Furthermore, if $\mathcal{M}$ is the collection of μ^*-measurable sets and $\overline{\mu} = \mu^*|_{\mathcal{M}}$, we have by the definition of μ^* that $\overline{\mu}$ is outer regular on all of $\mathcal{M}$. Also, $\overline{\mu}$ is fully determined by the linear functional $I(f) \coloneqq \int f d\mu$. (since I uniquely determines $\mu(U)$ for each open U).

Importantly, we can also say that $\mu^*(E) = \inf\{\mu(B) : B \supseteq E \text{ with } B \in \mathcal{B}_X\}$ for all $E \in \mathcal{P}(X)$. This is because (and I will be abusing notation to write this since otherwise it would be really cumbersome) if B represents any Borel set and U any open set, then:

$$\mu^*(E) \leq \mu^*(\bigcap_{B \supseteq E} B) \leq \inf_{B \supseteq E} \mu(B) \leq \inf_{U \supseteq E} \mu(U) = \mu^*(E).$$

Next, we claim that:

$$\inf\{\mu(B) : B \supseteq E \text{ with } B \in \mathcal{B}_X\} = \inf\{\sum_{n=1}^{\infty} \mu(B_n) : \text{ all } B_n \text{ are Borel and } E \subseteq \bigcup_{n \in \mathbb{N}} B_n\}$$

The fact that the right side is at most the left side is trivial. Meanwhile, to show the other inequality we can just note that if $(B_n)_{n \in \mathbb{N}}$ is a covering of E by Borel sets, then by taking differences we can get a sequence of disjoint Borel sets $(B'_n)_{n \in \mathbb{N}}$ covering E such that $B'_n \subseteq B_n$ for all n; $B := \bigcup_{n \in \mathbb{N}} B_n$ is a Borel set containing E; and:

$$\mu(B) = \sum_{n \in \mathbb{N}} \mu(B'_n) \leq \sum_{n \in \mathbb{N}} \mu(B_n)$$

As a result, μ^* is equal to the outer measure induced by treating μ as a premeasure on $(X, \mathcal{B}_X)$. Combining this with the previous exercise I did, we thus know that $(X, \mathcal{M}, \overline{\mu})$ is precisely the saturation of the completion of $(X, \mathcal{B}_X, \mu)$.

Now returning to Hewitt and Stromberg, here is one more lemma before I reset which variable names I have assigned to what.

Lemma 10.31: For any $A \subseteq X$, the following are equivalent:

(i) A is μ^*-measurable;

(ii) $\mu^*(U) \geq \mu^*(U \cap A) + \mu^*(U - A)$ for all open $U \subseteq X$ such that $\mu(U) < \infty$;

(iii) $A \cap U$ is μ^*-measurable for all open $U \subseteq X$ such that $\mu(U) < \infty$;

(iv) $A \cap K$ is μ^*-measurable for all compact $K \subseteq X$.

Proof:
It's trivial that (i) implies (iv).

Next suppose (iv) holds and let U be an open set such that $\mu(U) < \infty$. Then since μ is inner regular on all open sets, we know for each $n \in \mathbb{N}$ that there is a compact set $F_n \subseteq U$ such that $\mu(F_n) > \mu(U) - 1/n$. Letting $F = \bigcup_{n \in \mathbb{N}} F_n$ we have that $F \subseteq U$; F is μ^*-measurable (since all the F_n are); and $\mu(F) \geq \mu(F_n) > \mu(U) - 1/n$ for all n. It follows that $\mu(F) = \mu(U)$ and $\mu(U - F) = 0$.

One consequence of this is that:

$$\begin{aligned} A \cap U = A \cap (F \cup (U - F)) &= (A \cap F) \cup (A \cap (U - F)) \\ &= (\bigcup_{n \in \mathbb{N}} (A \cap F_n)) \cup (A \cap (U - F)). \end{aligned}$$

And since $A \cap (U - F) \subseteq U - F$ with $\mu(U - F) = 0$, we know by the completeness of $(X, \mathcal{M}, \overline{\mu})$ that $A \cap (U - F)$ is in $\mathcal{M}$. Also, since we have that all the $A \cap F_n$ are in $\mathcal{M}$ by (iv), we know that $A \cap U$ is a countable union of sets in $\mathcal{M}$. This proves (iii).

Now suppose (iii) and let $U \subseteq X$ be an open set such that $\mu(U) < \infty$. Then both U and $U \cap A$ are in $\mathcal{M}$ by (iii). And in turn we also have that $U - A = U - (U \cap A) \in \mathcal{M}$. Thus $\overline{\mu}(U) = \overline{\mu}(U \cap A) + \overline{\mu}(U - A)$ and we've shown (ii).

Finally suppose (ii) and let F be any subset of X. If $\mu^*(F) = \infty$, then we trivially have that $\mu^*(F) \geq \mu^*(A \cap F) + \mu^*(A - F)$. So assume $\mu^*(F) < \infty$. Then for any given $\varepsilon > 0$ we know there is an open set U such that $F \subseteq U$ and $\mu(U) < \mu^*(F) + \varepsilon$. In turn:

$$\mu^*(F) + \varepsilon > \mu(U) \geq \mu^*(U \cap A) + \mu^*(U - A) \geq \mu^*(F \cap A) + \mu^*(F - A)$$

Taking $\varepsilon \to 0$ finishes proving (i). ■

Oh, one more lemma I'll need is that $\overline{\mu}$ is inner regular on all of its σ-finite sets. (You can prove this identically to how we proved μ is inner regular on all of its σ-finite sets [see my math 240c notes]).

Theorem 19.30: Let X be an LCH space and let $(X, \mathcal{M}, \mu)$ be the saturation of the completion of a positive Borel Radon measure on X. Then there exists a family $\mathcal{F}_0$ of subsets of X with the following properties:

(i) the sets in $\mathcal{F}_0$ are compact and have finite measure greater than 0;

(ii) the sets in $\mathcal{F}_0$ are pairwise disjoint;

(iii) if $F \in \mathcal{F}_0$, U is open, and $U \cap F \neq \emptyset$, then $\mu(U \cap F) > 0$;

(iv) if $E \in \mathcal{M}$ and $\mu(E) < \infty$, then $\mu(E \cap F) > 0$ for only countably many $F \in \mathcal{F}_0$;

(v) the set $D := X - (\bigcup_{F \in \mathcal{F}} F)$ is measurable and locally null (which means that $\mu(E \cap A) = 0$ for all $A \in \mathcal{M}$ with $\mu(A) < \infty$ [see my homework from math 240b for more info about this...]);

(vi) if Y is a subset of X such that $Y \cap F \in \mathcal{M}$ for all $F \in \mathcal{F}_0$, then $Y \in \mathcal{M}$.

Proof:
Let $\mathcal{X}$ be the collection of all families of subsets in $\mathcal{M}$ which satisfy properties (i), (ii), and (iii). Then firstly note that $\emptyset \in \mathcal{X}$. So, $\mathcal{X}$ is not empty. Also, it easy to see that if $\mathcal{X}_0$ is a subset of $\mathcal{X}$ that is linearly ordered by the subset relation, then $\mathcal{X}_0$ has an upper bound in $\mathcal{X}$. Just set $\mathcal{F} = \bigcup_{\mathcal{F}'' \in \mathcal{X}_0} \mathcal{F}''$. Then it's obvious that $\mathcal{F}$ satisfies properties (i) and (iii) since every $F \in \mathcal{F}$ is in some $\mathcal{F}'' \in \mathcal{X}_0 \subseteq \mathcal{X}$. Also, if F_1 and F_2 are sets in $\mathcal{F}$, then we know there exists a collection $\mathcal{F}'' \in \mathcal{X}_0$ such that both $F_1, F_2 \in \mathcal{F}''$. And thus in turn we know that $F_1 \cap F_2 = \emptyset$, meaning $\mathcal{F}$ satisfies property (ii).

It follows by Zorn's lemma that $\mathcal{X}$ contains a maximal family $\mathcal{F}_0$. Our goal now is to show that $\mathcal{F}_0$ satisfies properties (iv), (v), and (vi).

Suppose $E \in \mathcal{M}$ and $\mu(E) < \infty$. Then for the sake of finding a contradiction, suppose $E \cap F \neq \emptyset$ for uncountably many $F \in \mathcal{F}_0$. Then by the outer regularity of μ, we know there exists an open set U such that $U \supseteq E$ and $\mu(U) < \mu(E) + 1 < \infty$. Then in turn, by property (ii) we have that $\mu(U) \geq \sum_{F \in \mathcal{F}_0} \mu(U \cap F)$ (because U is a superset of any finite union of the $F \cap U$). But now since $U \cap F \supseteq E \cap F \neq \emptyset$ for uncountably many $F \in \mathcal{F}_0$, we know by property (iii) that $\mu(U \cap F) > 0$ for uncountably many $F \in \mathcal{F}_0$. But that implies that $\sum_{F \in \mathcal{F}_0} \mu(U \cap F) = \infty$, which is a contradiction. Hence, we've proven property (iv).

Next let U be any open set with $\mu(U) < \infty$ and let $\mathcal{F}_1$ be the countable subfamily of $\mathcal{F}_0$ containing all the F such that $\mu(U \cap F) > 0$. Since all $F \in \mathcal{F}_1$ are in $\mathcal{M}$, we know that $\bigcup_{F \in \mathcal{F}_1} F \in \mathcal{M}$. Thus $\mu(U) = \mu(U \cap \bigcup_{F \in \mathcal{F}_1} F) + \mu(U - \bigcup_{F \in \mathcal{F}_1} F)$. But by (iii) we have that $\mathcal{F}_1 = \{F \in \mathcal{F}_0 : U \cap F \neq \emptyset\}$. Thus, it's clear that $U \cap \bigcup_{F \in \mathcal{F}_1} F = U \cap \bigcup_{F \in \mathcal{F}_0} F$ and $U - \bigcup_{F \in \mathcal{F}_1} F = U - \bigcup_{F \in \mathcal{F}_0} F$. And by the previous lemma, this proves that $\bigcup_{F \in \mathcal{F}_0} F$ and $D \coloneqq X - \bigcup_{F \in \mathcal{F}_0} F$ are in $\mathcal{M}$.

Now we still need to show that D is locally null. So suppose for the sake of contradiction that there exists $A \in \mathcal{M}$ with $0 < \mu(A \cap D) < \infty$. Then since μ is inner regular on $A \cap D$, there'd be a compact set $K \subseteq A \cap D$ such that $0 < \mu(K) < \mu(A \cap D) < \infty$. Also, if we consider the collection $\mathcal{U}$ of open sets $U \subseteq X$ such that $\mu(U \cap K) = 0$, then $K \cap \bigcup_{U \in \mathcal{U}} U$ is measurable on account of $\bigcup_{U \in \mathcal{U}} U$ being open. We claim $K \cap \bigcup_{U \in \mathcal{U}} U$ is a null set.

> Otherwise, by inner regularity there would exist a compact set $C \subseteq K \cap \bigcup_{U \in \mathcal{U}} U$ with $\mu(C) > 0$. And since $\mathcal{U}$ is an open cover of C, we'd have that there is a finite subcover of sets $U \cap H$, all with measure zero, covering C. This is a contradiction.

Setting $H \coloneqq K - \bigcup_{U \in \mathcal{U}} U$, we'd know that H is compact on account of being a closed subset of K. Also, we'd know that $\mu(H) = \mu(K) > 0$. And it's clear that H would be disjoint from all the $F \in \mathcal{F}_0$. And finally, if V were any open set such that $V \cap H \neq \emptyset$ then we'd know that $V \notin \mathcal{U}$. But that would mean that $\mu(V \cap K) > 0$. And so:

$$\mu(V \cap H) = \mu(V \cap K) - \mu(V \cap (K - \textstyle\bigcup_{U \in \mathcal{U}} U)) = \mu(V \cap K) - 0 > 0$$

Hence, we've shown that $\mathcal{F}_0 \cup \{H\}$ is a collection in $\mathcal{X}$ that is strictly larger than $\mathcal{F}_0$. But that contradicts that $\mathcal{F}_0$ is maximal. Thus, we've proven property (v).

Finally, let Y be as in the theorem statement and consider any open set $U \subseteq X$ such that $\mu(U) < \infty$. Then:

$$U \cap Y = (U \cap Y \cap D) \cup (U \cap Y \cap \bigcup_{F \in \mathcal{F}_0} F) = (U \cap Y \cap D) \cup \bigcup_{F \in \mathcal{F}_0} (U \cap (Y \cap F))$$

But now we know from before that there is a countable subfamily of $\mathcal{F}_0$ containing all the F which intersect U. Also, since D is locally null, we know that $\mu(U \cap D) = 0$ and in turn $U \cap Y \cap D$ is measurable due to it being a subset of a null set. It follows that $U \cap Y$ is a countable union of measurable sets. So, $U \cap Y$ is measurable. By our prior lemma, we thus have that $Y \in \mathcal{M}$. This proves (vi). ■

> Side note: You may note that since D is locally null and all compact sets have finite measure, we must have that any compact set $K \subseteq X$ contained in D has measure zero. By inner regularity, this in turn implies that any open set $U \subseteq X$ contained in D has measure zero.

<u>Corollary 19.31:</u> Let X be an LCH space and let $(X, \mathcal{M}, \mu)$ be the saturation of the completion of a positive Borel Radon measure on X. Then $(X, \mathcal{M}, \mu)$ is decomposable.

Proof:

Let $\mathcal{F}_0$ and D be as in the last theorem. Then set $\mathcal{F} \coloneqq \mathcal{F}_0 \cup \{\{x\} : x \in D\}$. We claim $\mathcal{F}$ is a decomposition.

(i) Since X is Hausdorff, we know that $\{x\}$ is compact for all $x \in D$. Hence, every set in $\mathcal{F}$ is compact and thus has finite measure.

(ii) It's clear that all the elements of $\mathcal{F}$ form a partition of X.

(iii) If $\mu(E) < \infty$ then since D is locally null, we have that $\mu(E \cap D) = 0$. Also, by how we chose $\mathcal{F}_0$ we know there are only countably many $F \in \mathcal{F}_0$ with $F \cap E \neq \emptyset$. So:

$$\mu(E) = \sum_{F \in \mathcal{F}} \mu(E \cap F).$$

(iv) This is an immediate consequence of the sixth property that we proved about $\mathcal{F}_0$. ■

Corollary 19.32: Another Extension of the Lebesgue-Radon-Nikodym Theorem:

Let X be an LCH space and let $(X, \mathcal{M}, \mu)$ be the saturation of the completion of a positive Borel Radon measure on X. Also let ν be any measure on $(X, \mathcal{M})$ such that $\nu \ll \mu$. Then all the conclusions of theorem 19.27 (on page 156 of my journal) hold.

Reflection:

It is here that Hewitt and Stromberg mostly stop being useful. So, I will give one last quote from them before expositing some more about this topic on my own (also geeze I just noticed that this quote has a typo. It's supposed to be "(19.32)"; not "(19.33)"...):

> [redacted] One may well ask if the generality obtained in (19.27) and (19.33) is worth the effort. Many mathematicians believe it is not. But we feel it our duty to show the reader the most general theorems that we can reasonably produce and that he might reasonably need. [redacted]

I'm not gonna lie, while going down research rabbitholes is really fun, I feel like Folland really oversold the significance of this result. When I went into this rabbithole, I was under the impression from Folland that I'd be proving that if μ is an arbitrary Radon measure and $\nu \ll \mu$, then ν always has the form $f d\mu$. But that is not what was proven in theorems 19.27 and 19.32 so I don't know what Folland was yappin about. Considering the lack of specific details that Folland gives, I can't help but suspect that Folland didn't entirely understand what Hewitt and Stromberg actually proved. Although, I don't necessarily blame Folland for that.

Hewitt and Stromberg were difficult for me to understand partly because they had a lot of (in my opinion) very questionable conventions. For example, I tried my best to hide it away as much as possible but Hewitt and Stromberg explicitely treated every measure as the restriction of an outer measure to the σ-algebra of μ^*-measurable sets rather than abstracting that away because I guess they always wanted their σ-algebras to be as large as possible. In fact, while I was skimming the book I even saw some other sections where they talked about extending their measures to even larger σ-algebras. Never mind that there are advantages to working on a smaller σ-algebra (see for instance page 55 of my journal: when working on a larger σ-algebra we lose that continuity implies measurability...).

I'll also mention that the choice of math font in their book is awful and I couldn't help but wonder if they knew what kerning is. Anyways I guess my review of their book is that it feels dated and very clunky. Although, maybe I only had a mixed experience because I'm not an intended reader. (cough cough)

Anyways, while getting a milkshake with my apartment-mates, I thought of a few ideas for how to make the preceding theorems actually useful.

My first challenge is that I'm primarily working with Borel Radon measures while Hewitt and Stromberg weren't. This leads me to the following result:

Proposition: Suppose X is an LCH space and let μ and ν be positive Radon measures on $(X, \mathcal{B}_X)$ such that $\nu \ll \mu$. Then let $(X, \mathcal{M}_\mu, \overline{\mu})$ be the saturation of the completion of $(X, \mathcal{B}_x, \mu)$ and let $(X, \mathcal{M}_\nu, \overline{\nu})$ be the saturation of the completion of $(X, \mathcal{B}_x, \nu)$.

- In general $\mathcal{M}_\mu \neq \mathcal{M}_\nu$.

  Proof:
  Let $X = \mathbb{R}$. Then set μ to be the Lebesgue measure and set ν equal to the zero measure (i.e. the measure for which every set is null). Note that we trivially have that $\nu \ll \mu$. Also, both are easily seen to be Radon measures.

  Since μ and ν are σ-finite, we know by the exercise on pages 158 and 159 that $\mathcal{M}_\mu$ and $\mathcal{M}_\nu$ are just the completions of $\mathcal{M}$ with respect to μ and ν respectively. Since $\nu(X) = 0$, we have that $\mathcal{M}_\nu = \mathcal{P}(X)$. Meanwhile, because of the existance of Vitali sets, we know that $\mathcal{M}_\mu \neq \mathcal{P}(X)$. ■.

- We do always have that $\mathcal{M}_\mu \subseteq \mathcal{M}_\nu$.

  Proof: (I got this argument from Hewitt and Stromberg 19.33)
  Suppose $A \in \mathcal{M}_\mu$ and let $F \subseteq X$ be any compact set. Since A is locally measurable with respect to the completion of μ, we know that $A \cap F$ is in the completion of $\mathcal{B}_X$ with respect to μ. So let $A \cap F = E \cup N$ where $E \in \mathcal{B}_X$ and $N \subseteq N'$ with $N' \in \mathcal{B}_X$ and $\mu(N') = 0$. Then since $\mu(N') = 0$ implies that $\nu(N') = 0$, we know that $A \cap F$ is in the complection of $\mathcal{B}_X$ with respect to ν. By the lemma on page 161, this proves that $A \in \mathcal{M}_\nu$. ■

- We do always have that $\overline{\nu}|_{\mathcal{M}_\mu} \ll \overline{\mu}$.

  Proof:
  Suppose $A \in \mathcal{M}_\mu$ with $\overline{\mu}(A) = 0$. Since A has finite measure, we know that A is not merely locally measurable but also that A is in the completion of $\mathcal{B}_X$ with respect to μ. Hence, there exists a set $N \in \mathcal{B}_X$ such that $A \subseteq N$ and $\mu(N) = 0$. But now since $\nu \ll \mu$, we have that $\nu(N) = 0$. In turn, $\overline{\nu}(A) \leq \nu(N) = 0$. This proves that $\overline{\mu}(A) = 0 \Longrightarrow \overline{\nu}(A) = 0$. ■

Bibliography:

Everything is cited in the order it shows up in the journal with the exception that the first citation is for the book that got me to actually sit down and make a bibliography. Also, I decided to write my citations according to APA 7.

1. Hewitt, E., & Stromberg, K. (1975). *Real and Abstract Analysis*. Springer.
2. Rudin, W. (1976). *Principles of Mathematical Analysis* (3rd ed.). McGraw Hill.
3. Munkres, J. R. (2014). *Topology*. Pearson.
4. Artin, M. (2010). *Algebra* (2nd ed.). Pearson.
5. Folland, G. B. (1999). *Real Analysis : Modern Techniques and Their Applications* (2nd ed.). John Wiley And Sons.
6. Munkres, J. R. (1994). *Analysis on manifolds*. Addison-Wesley.
7. Guillemin, V., & Haine, P. (2019). *Differential Forms*. World Scientific Publishing Company.
8. Spencer, J. (1981). Balancing unit vectors. *Journal of Combinatorial Theory, Series A*, 30(3), 349–350. https://doi.org/10.1016/0097-3165(81)90033-9
9. Lee, J. M. (2013). *Introduction to smooth manifolds*. Springer.