# Math Journal

Isabelle Mills

July 13, 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 $\alpha$ 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 $\gamma = \bigcup\limits_{\alpha \in A} \alpha$. 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 $\gamma$ 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 $\gamma$ 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, $\gamma$ satisfies property 3. of cuts.

With that, we've now shown that $\gamma \in R$. Clearly, $\gamma$ 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 $\delta$ 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 $\alpha \in R$, $\alpha + 0^* = \alpha$.

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 $\alpha$. Secondly, we add the $-r$ term to guarentee that $\beta$ 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 $\beta$ is the additive inverse of $\alpha$. 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 $r \in \alpha$, $s \in \beta$, $r > 0$, 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 $\alpha$ and $\beta$ 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 $\beta$ 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 $X^\omega$.

For any set $A \in \mathcal{P}(\mathbb{Z}_+)$, define $f(A)$ to be the $\omega$-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 $X^\omega$.

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 $F = \bigcup_{n=1}^{\infty} A_n$.

(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 $G = \bigcup\limits_{n=1}^{\infty} A_n$.

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, $H = \bigcup\limits_{n=1}^{\infty} A_n$.

(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 $\bigcup\limits_{n=1}^{\infty} (\mathbb{Z}_+)^n$ 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, $S_1 = \bigcup_{n=1}^{\infty}(A_n - B_n)$ 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 $\rho$ 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 $\alpha$ of $S^{\mathsf{C}}$.

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 $\beta$ 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 $\bigcup_{A \in \mathcal{A}} A \times \bigcup_{A \in \mathcal{A}} A$.

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 $C = \{a \in \bigcup_{A \in \mathcal{A}} A \mid \exists A \in \mathcal{A} \; s.t. \; a \in A$ 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 $\beta$. This $\beta$ 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 $S_\alpha$ denote the set $\{x \in X \mid x < \alpha\}$. We call $S_\alpha$ the section of $X$ by $\alpha$.

**Lemma 10.2:** There exists a well-ordered set $A$ having a largest element $\Omega$ such that $S_\Omega$ 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 $\Omega$ be the least element of $C$ such that $S_\Omega$ 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 $S_\Omega$, then $A$ has an upper bound in $S_\Omega$.

Proof:
Let $A$ be a countable subset of $S_\Omega$. For all $a \in A$, we know that $S_a$ is countable. Therefore, $B = \bigcup_{a \in A} S_a$ 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 $S_\Omega$ be a minimal uncountable well-ordered set.

(a) Show that $S_\Omega$ has no largest element.

Suppose $\alpha \in S_\Omega$ is the largest element of $S_\Omega$. In that case, we'd have that $S_\alpha = S_\Omega - \{\alpha\}$. However, by theorem 10.3, we know that $S_\alpha$ 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 $S_\Omega$ 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: $x \not< x$<br>Transitivity: $x < y$ and $y < z \Rightarrow x < z$<br>Comparability: $x < y$ or $y < x$ is true | Nonreflexivity: $x \not\prec x$<br>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 $\alpha$ 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 $\alpha$ has $\beta$ as its successor. Meanwhile, if $S \neq \emptyset$, then since $A$ is well-ordered, we know that $A$ has a least element $\gamma$. Thus, the set $\{x \in A \mid \alpha < x < \gamma\} = \emptyset$ and we know that $\gamma$ is the successor of $\alpha$.

(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 $S_\Omega$ 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 $S_\Omega$ has no largest element according to exercise 10.6.a. So, all elements of $S_\Omega$ 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 $S_\Omega$ and hence $T$ is well-ordered, we know that $R$ has a least element $\beta$. But note that $\beta$ has a predecessor $\gamma$ which isn't in $R$. More specifically, since we know that the least element of $T$ is in $f(\mathbb{Z}_+)$, we know that $\gamma$ 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 $\beta$ 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 $S_\Omega$.

**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 $(S_\alpha \subseteq J_0) \Longrightarrow \alpha \in J_0$.

**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 $\alpha$ be the least element of $J - J_0$. We know that $S_\alpha$ 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 $\alpha$ 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 $k(\alpha) = \text{smallest}(C - h(S_\alpha))$.

(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 $k : \bigcup_{\alpha \in K} S_\alpha \longrightarrow C$ satisfying $(*)$.

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

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 $x \in \bigcup_{\alpha \in K} S_\alpha$, 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 $h_\alpha$ 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: $\beta$ has an immediate predecessor $\alpha$.

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

Case 2: $\beta$ has no immediate predecessor.

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

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 $\bigcup_{\alpha \in S_\beta} S_\alpha = S_\beta$ 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 $\beta$.

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 $a_\alpha$ 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 $a_\alpha$ is comparable to $a_\beta$ 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 $a_\alpha$ 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 $a_\gamma$ is not comparable to $a_\alpha$. 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 $\mu$ 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 $\infty$.

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

**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 $b \preceq a \Longrightarrow a = b$.

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 $a \prec b \prec a \Longrightarrow a \prec a$ which is not allowed.

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

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 $\mathcal{A}$: $\bigcup_{B \in \mathcal{B}} B$.

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.

$$
\begin{tikzcd}
 & \text{Non-measurable Sets} \\
\text{Axiom of Choice} \arrow[ur, Rightarrow, blue] \arrow[d, Rightarrow, shift left, red] & \\
\text{Well Ordering Theorem} \arrow[u, Rightarrow, shift left, blue] \arrow[d, Rightarrow, shift left, blue] & \\
\text{Maximum Principle} \arrow[u, Rightarrow, shift left, red] \arrow[d, Rightarrow, blue] & \text{Tukey's Lemma} \arrow[l, Rightarrow, blue] \\
\text{Zorn's Lemma} \arrow[r, Rightarrow, blue] & \text{Kuratowski's Lemma} \arrow[u, Rightarrow, blue]
\end{tikzcd}
$$

**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: $\alpha$ has an immediate predecessor $\beta$.

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: $\alpha$ has no immediate predecessor.

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

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 $\alpha$.

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 $h(J) = h(\bigcup\limits_{\alpha \in J} S_\alpha) = \bigcup\limits_{\alpha \in J} h(S_\alpha) = \bigcup\limits_{\alpha \in J} S_{h(\alpha)}$.

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 $S_\beta$ and $B$ has the order type of $S_\alpha$ 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 $S_\beta$ and $g$ is the order-preserving bijection from $B$ to $S_\alpha$.

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 $S_\alpha$. This gives us a contradiction since exercise 3 would then imply that $A$ has the same order type as either $S_\alpha$ or a section of $S_\alpha$ (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 $\beta \in B$, $S_\alpha$ and $S_{(A,\alpha)}$ have the same order type, as do $S_\beta$ 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 $S_\alpha$ 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 $\alpha$ 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 $\beta$ 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 $\alpha$ be the least element (using $<_1$) of $B_1 \cap S$.

We claim that $\alpha$ 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 $\alpha$ 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 $\alpha$.
- 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 $\alpha$. 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 $U_\alpha$ 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 $U = \bigcup\limits_{x \in U} B_x$, 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 $\alpha$, we know that $\bigcup\limits_{\beta \in B} U_\beta \in \mathcal{T}_\alpha$ for all $\alpha$. 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 $U_\beta$ 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:
  $$abb^{-1}a^{-1} = a1a^{-1} = aa^{-1} = 1 \text{ and } b^{-1}a^{-1}ab = b^{-1}1b = b^{-1}b = 1$$

---

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 $1, 2, \ldots n$.

---

$S_3$ 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 $G(x) := \boldsymbol{E}(F(x)) = \exp(F(x))$ where $\exp(x) := \sum_{n\geq 0} \frac{x^n}{n!}$. 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 $n$th. 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 $m$th. 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 $m$th. root of $G(x)$ with $1$ as its constant coefficient.

Because of part (f), we can now extend our definition of the $z$th. 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:
  $$\begin{aligned}\mathrm{D}H(x) &= \mathrm{D}(\boldsymbol{E}(z \cdot \boldsymbol{L}(G(x))))(x)\\ &= \mathrm{D}(z \cdot \boldsymbol{L}(G(x)))(x) \cdot \boldsymbol{E}(z \cdot \boldsymbol{L}(G(x)))\\ &= z \cdot \mathrm{D}(\boldsymbol{L}(G(x)))(x) \cdot G(x)^z = z\frac{\mathrm{D}G(x)}{G(x)}G(x)^z\\ &= z\mathrm{D}G(x)G(x)^{-1}G(x)^z\\ &= z\mathrm{D}G(x)G(x)^{z-1}\end{aligned}$$

- Binomial Theorem: Given $z \in \mathbb{C}$, we have that:
  $(1+x)^z = \sum\limits_{n\geq 0} \binom{z}{n} x^n$ 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:

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

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

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 $\rho$ 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 $\alpha$ times the length of $I$ the "open middle $\alpha$th" 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 $\alpha_j$th from the intervals that make up $K_{j-1}$. Then $K = \bigcap\limits_{j\in\mathbb{N}} K_j$ is called a generalized Cantor set.

The ordinary Cantor set $C$ is obtained by setting all $\alpha_j$ 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 $\alpha_j$. 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 $j$th 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 $0$s and $1$s 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:

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

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 $\beta$. 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 $g(x)$ to all $\mathbb{R}$ (One way to do this would be to set $G(x) = 2x$ when $x \notin [0,1]$). Then set $F = \chi_B$ where $B$ is the the set found in part $c$. Then $(F \circ G)^{-1}(\{1\}) = A$ is not Lebesgue measurable. So $F \circ G$ is not a Lebesgue measurable function.

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 $S_\Omega$ be a minimal uncountable set (by constructing $S_\Omega$ 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 $S_\Omega$, 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 $\alpha$ has a direct predecessor $\beta$, then let $\mathcal{E}_\alpha$ be the collection of all countable unions of and complements of sets from $\mathcal{E}_\beta$.
- If $\alpha$ 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: The Dual of $C_0(X)$. 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 $\mu$, we need to determine when $I(f) = \int f \mathrm{d}\mu$ is bounded.

Since $X$ is open and $\mu$ 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 $C_0(X)$.

Lemma 7.15: 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 $c = 0$:

$$\begin{aligned}I^+(cf) &= \operatorname{sgn}(c)\left(I^+(|c|f^+) - I^+(|c|f^-)\right)\\ &= \operatorname{sgn}(c)|c|\left(I^+(f^+) - I^+(f^-)\right) = cI^+(f).\end{aligned}$$

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:

$$\begin{aligned}I^+(f^+) + I^+(g^-) + I^+(h^-) &= I^+(f^+ + g^- + h^-)\\ &= I^+(f^- + g^+ + h^+) = I^+(f^-) + I^+(g^+) + I^+(h^+)\end{aligned}$$

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

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 $J_1(f) = \operatorname{Re}(J(f))$ and $J_2 = \operatorname{Im}(J(f))$. 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 $I_1^+(f) = J_1^+(\operatorname{Re}(f)) + iJ_1^+(\operatorname{Im}(f))$ 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:

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

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:

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

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 $\sigma$-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 $\mu$.

Proposition 7.16: If $\mu$ is a complex Borel measure, then $\mu$ 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 $\sigma$-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 $\mu_j$ 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 $\nu$ 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 $\min(f_3, f_4) = 0$ $\nu$-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 $\nu$, 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 $\mu$ respectively, it's not necessarily clear that $\mu_1$ and $\mu_2$ are the positive and negative variations respectively of $(\mu_1 - \mu_2)$ and likewise for $\mu_3$ and $\mu_4$ 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.

Lemma (Riesz Representation theorem on $C_0(X,\mathbb{R})^*$): 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 $\mu$ 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 $\mu$, $f \mapsto \int f\mathrm{d}\mu$ 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 $\mu_1$ and $\mu_2$ 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 $I(f) = \int f\mathrm{d}\mu$.

Exercise 7.16: The positive and negative variations of $\mu$ are the Radon measures $\mu_1$ and $\mu_2$ respectively.

Let $\mu^+$ and $\mu^-$ be the positive and negative variations of $\mu$, 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 $\mu_1, \mu_2$ 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, $\mu$ 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 $\mu^+$ 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 $\mu$ 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 $I_\mu$ for some $\mu \in M(X)$. On the other hand, if $\mu \in M(X)$, then we already know that $I_\mu$ is a linear function. Also, by proposition 3.13:

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

So, $I_\mu$ 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:

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

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:

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

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:

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

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 $\chi^\wedge_{[-a,a]} = \chi^\vee_{[-a,a]} = 2a\frac{\sin(2a\pi x)}{2\pi a x} = \frac{\sin(2\pi a x)}{\pi x}$.

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:

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

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

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

(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, $|\frac{\sin(y)}{y}|^2 \leq \chi_{[-1,1]}(y) + \frac{1}{y^2}\chi_{[-1,1]^\mathsf{C}}(y)$ 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 $\chi^{\vee}_{[-a,a]} = \frac{\sin(2\pi a x)}{\pi x}$ 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:

$$\pi\int_{-\frac{1}{2\pi}}^{\frac{1}{2\pi}} \frac{s}{s^2+4\pi^2\xi^2}\mathrm{d}\xi = \frac{s\pi}{4\pi^2}\int_{-\frac{1}{2\pi}}^{\frac{1}{2\pi}} \frac{1}{(\frac{s}{2\pi})^2+\xi^2}\mathrm{d}\xi$$

$$= \frac{s}{4\pi}\left(\frac{2\pi}{s}\right)\left[\arctan(\tfrac{2\pi}{s}\xi)\right]_{\xi=-\frac{1}{2\pi}}^{\xi=\frac{1}{2\pi}} = \frac{1}{2}\left(\arctan(\tfrac{1}{s}) - \arctan(-\tfrac{1}{s})\right) = \arctan(\tfrac{1}{s})$$

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:

$$\int_0^b \frac{\sin(y)}{y}e^{-sy}\mathrm{d}y = \sum_{j=0}^{n-1}(-1)^j\int_{j\pi}^{(j+1)\pi}|\frac{\sin(y)}{y}e^{-sy}|\mathrm{d}y + \int_{n\pi}^{n\pi+\alpha}|\frac{\sin(y)}{y}e^{-sy}|\mathrm{d}y$$

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 $c_n = \int_{j\pi}^{(j+1)\pi}|\frac{\sin(y)}{y}e^{-sy}|\mathrm{d}y$, 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 $\varepsilon$ 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:

$$\begin{aligned}\left|\int_0^\infty \frac{\sin(y)}{y}\mathrm{d}y - \frac{\pi}{2}\right| &\leq \left|\int_0^\infty \frac{\sin(y)}{y}\mathrm{d}y - \int_0^b \frac{\sin(y)}{y}\mathrm{d}y\right| + \left|\int_0^b \frac{\sin(y)}{y}\mathrm{d}y - \int_0^b \frac{\sin(y)}{y}e^{-sy}\mathrm{d}y\right| \\ &\quad + \left|\int_0^b \frac{\sin(y)}{y}e^{-sy}\mathrm{d}y - \int_0^\infty \frac{\sin(y)}{y}e^{-sy}\mathrm{d}y\right| + \left|\int_0^\infty \frac{\sin(y)}{y}e^{-sy} - \frac{\pi}{2}\right| \\ &< \varepsilon/4 + \varepsilon/4 + \varepsilon/4 + \varepsilon/4 = \varepsilon\end{aligned}$$

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:

$\lim\limits_{t\to 0} \frac{(f*g)(x+te_j)-(f*g)(x)}{t} = ((\partial_j f) * g)(x)$ iff $\lim\limits_{t\to 0} \left|\frac{(f*g)(x+te_j)-(f*g)(x)}{t} - ((\partial_j f) * g)(x)\right| = 0.$

Now for all $t \neq 0$:

$$\begin{aligned}0 &\leq \left|\tfrac{(f*g)(x+te_j)-(f*g)(x)}{t} - ((\partial_j f)*g)(x)\right| \\ &= \left|t^{-1}\int (f(x+te_j-y)-f(x-y))\,g(y)\mathrm{d}y - \int \partial_j f(x-y)g(y)\mathrm{d}y\right| \\ &= \left|\int t^{-1}(\tau_{-te_j}f - f)(x-y)g(y)\mathrm{d}y - \int \partial_j f(x-y)g(y)\mathrm{d}y\right| \\ &\leq \int |t^{-1}(\tau_{-te_j}f-f)(x-y) - \partial_j f(x-y)||g(y)|\mathrm{d}y \\ &\leq \|t^{-1}(\tau_{-te_j}f-f)-\partial_j f\|_p \|g\|_q\end{aligned}$$

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, $\varphi$ 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 $(L_g)^{\mathsf{C}}$ 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:

$$\begin{aligned}|h*\varphi_{1/n}(x)| \leq \int |h(x-y)\varphi_{1/n}(y)|\mathrm{d}y &\leq \|h\|_p \left(\int_{-1/n}^{1/n} |n(1-|nx|)|^q \mathrm{d}x\right)^{1/q} \\ &\leq n\|h\|_p \left(\int_{-1/n}^{1/n} 1\mathrm{d}x\right)^{1/q} = n\left(\tfrac{2}{n}\right)^{1/q}\|h\|_p \leq 2^{1/q}\|h\|_p\end{aligned}$$

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:

$f(x) = \int_{a_1}^x h(y)\mathrm{d}y + f(a_1) = \int_{a_2}^x h(y)\mathrm{d}y + f(a_2)$ 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, $\mathcal{F}(\tau_{-y}f) = \lim\limits_{n\to\infty} \mathcal{F}(\tau_{-y}f_n)$.

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, $\mathcal{F}(\tau_{-y}f) = \lim\limits_{n\to\infty} e^{2\pi i\xi y}\widehat{f_n}(\xi) = e^{2\pi i\xi y}\widehat{f}(\xi)$ 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 $2\pi$ 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 singular 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 $\mathrm{NBV}$. After all, they are in $\mathrm{BV}$ 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 $\mu_{H_k}$ and$\mu_H$ 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 $\mu_{H_k}$ and $\mu_H$ with respect to the Lebesgue measure. Then since $\mu_H$ and $\mu_{H_k}$ are both finite measures in the separable locally compact Hausdorff space $\mathbb{R}$, we know by theorem 7.8 that $\mu_{H_k}$ and $\mu_H$ 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 $\beta$ 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 uncertainity 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 $\beta$, and that will be small precisely when the probability distribution of the property has a small standard deviation and $\beta$ 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:

$$\begin{aligned} \int_E |\widehat{f}|^2 &= \int \chi_E(\xi) \left| \int f(x) e^{-2\pi i \xi \cdot x} \mathrm{d}x \right|^2 \mathrm{d}\xi \\ &= \int \left| \int f(x) e^{-2\pi i \xi \cdot x} \sqrt{\chi_E(\xi)} \mathrm{d}x \right|^2 \mathrm{d}\xi \\ &\leq \left[ \int \left( \int |f(x) e^{-2\pi i \xi \cdot x} \sqrt{\chi_E(\xi)}|^2 \mathrm{d}\xi \right)^{1/2} \mathrm{d}x \right]^2 = \left[ \int |f(x)| \left( \int_E \mathrm{d}\xi \right)^{1/2} \mathrm{d}x \right]^2 = m(E) \left( \int |f(x)| \mathrm{d}x \right)^2 \end{aligned}$$

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 $\int_E |\widehat{f}|^2 \leq m(E)(\sqrt{m(S)}\|f\|_2)^2 = m(E)m(S)\|f\|_2^2$. ■