Math 140B Lecture Notes (Professor: Brandon Seward)

Isabelle Mills

May 18, 2024

Lecture 1: 4/1/2024

Let $f : E \longrightarrow \mathbb{R}$ where $E \subseteq \mathbb{R}$.

Since E is the domain of f, we shall also refer to it as $\operatorname{dom}(f)$.

Fix a point $x \in E \cap E'$. Then consider the function $\frac{f(t)-f(x)}{t-x}$ for $t \in \operatorname{dom}(f) \setminus \{x\}$ and define the derivative of f at x to be $f'(x) = \lim_{t \to x} \left(\frac{f(t)-f(x)}{t-x} \right)$ provided that this limit exists. When the above limit exists, we say f is differentiable at x.

We say f is differentiable on $D \subseteq E$ if f is differentiable at every point in D, and if f is differentiable on its entire domain, then we call f differentiable.

The function $f'(x) = \lim_{t \to x} \left(\frac{f(t)-f(x)}{t-x} \right)$ is called the derivative of f.

Proposition 83: If f is differentiable at x, then f is continuous at x.

Proof:
Note that $\lim_{t \to x} (f(t)) = \lim_{t \to x} \left((t-x) \frac{f(t)-f(x)}{t-x} + f(x) \right)$.

Now $\lim_{t \to x} (t-x) = 0$ and we know $\lim_{t \to x} \frac{f(t)-f(x)}{t-x} = f'(x)$ exists because f is differentiable at x. Also, obviously $\lim_{t \to x} f(x) = f(x)$.

Thus by proposition 66 (check 140A notes), we know that:

$$\begin{aligned}
\lim_{t \to x} \left((t-x) \frac{f(t)-f(x)}{t-x} + f(x) \right) &= \lim_{t \to x} (t-x) \lim_{t \to x} \left(\frac{f(t)-f(x)}{t-x} \right) + \lim_{t \to x} f(x) \\
&= 0 \cdot f'(x) + f(x) \\
&= f(x)
\end{aligned}$$

Thus, f is continuous at x.

Notes:

1. The above proposition says that differentiability is stronger than continuity.
2. The converse of this proposition is false. For example, the function $f(x) = |x|$ is continuous at $x = 0$ but not differentiable at $x = 0$.

Proposition 84: Suppose f and g are real-valued functions with $\mathrm{dom}(f), \mathrm{dom}(g) \subseteq \mathbb{R}$. Also suppose f and g are differentiable at x. Then $f+g$, fg, and (when $g(x) \neq 0$) $\frac{f}{g}$ are differentiable at x with:

(A) $$(f+g)'(x) = f'(x) + g'(x)$$ (sum rule)

(B) $$(fg)'(x) = f'(x)g(x) + f(x)g'(x)$$ (product rule)

(C) $$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}$$ (quotient rule)

Proof:

(A) Since both f and g are differentiable, we know that both $f'(x) = \lim\limits_{t\to x} \frac{f(t)-f(x)}{t-x}$ and $g'(x) = \lim\limits_{t\to x} \frac{g(t)-g(x)}{t-x}$ exist. So by proposition 66:

$$(f+g)'(x) = \lim_{t\to x} \frac{f(t)+g(t)-f(x)-g(x)}{t-x} = \lim_{t\to x} \frac{f(t)-f(x)}{t-x} + \lim_{t\to x} \frac{g(t)-g(x)}{t-x}$$

This means that $(f+g)'(x) = f'(x) + g'(x)$.

(B) Note that:

$$\begin{aligned}(fg)'(x) &= \lim_{t\to x} \frac{f(t)g(t)-f(x)g(x)}{t-x} \\ &= \lim_{t\to x} \frac{f(t)g(t)-f(x)g(t)+f(x)g(t)-f(x)g(x)}{t-x} \\ &= \lim_{t\to x} \left(g(t)\frac{f(t)-f(x)}{t-x} + f(x)\frac{g(t)-g(x)}{t-x}\right)\end{aligned}$$

By proposition 83, $g(t) \to g(x)$ as $t \to x$. Also, since both f and g are differentiable, we know $f'(x) = \lim\limits_{t\to x} \frac{f(t)-f(x)}{t-x}$ and $g'(x) = \lim\limits_{t\to x} \frac{g(t)-g(x)}{t-x}$ exist. So by proposition 66:

$$\lim_{t\to x} \left(g(t)\frac{f(t)-f(x)}{t-x} + f(x)\frac{g(t)-g(x)}{t-x}\right) = f'(x)g(x) + f(x)g'(x).$$

(C) Note that:

$$\begin{aligned}\left(\tfrac{f}{g}\right)'(x) &= \lim_{t\to x} \frac{\frac{f(t)}{g(t)} - \frac{f(x)}{g(x)}}{t-x} \\ &= \lim_{t\to x} \left(\frac{1}{g(x)g(t)} \frac{f(t)g(x)-f(x)g(t)}{t-x}\right) \\ &= \lim_{t\to x} \left(\frac{1}{g(x)g(t)} \frac{f(t)g(x)-f(x)g(x)+f(x)g(x)-f(x)g(t)}{t-x}\right) \\ &= \lim_{t\to x} \left(\frac{1}{g(x)g(t)} \left(g(x)\frac{f(t)-f(x)}{t-x} - f(x)\frac{g(t)-g(x)}{t-x}\right)\right)\end{aligned}$$

Now, for the same reasons as before, we can use propositions 83 and 66 to separate the parts of the above limit to get that the above limit equals:

$$\frac{1}{(g(x))^2}\left(g(x)f'(x) - f(x)g'(x)\right)$$

If $f(x) = \alpha$ where $\alpha \in \mathbb{R}$ is constant, then trivially $f'(x) = 0$ for all x.
Meanwhile, if $f(x) = x$, then we can trivially find that $f'(x) = 1$.

Claim 1: For all $n \in \mathbb{Z}_+$, if $f(x) = x^n$, then $f'(x) = nx^{n-1}$.

Proof: (we proceed by induction)

Base Case:

If $n = 1$, then for $f(x) = x^1$, we have that $f'(x) = 1 \cdot x^0$.

Induction:

Now assume $n > 1$, and for $f(x) = x^{n-1}$, we have that $f'(x) = (n-1)x^{n-2}$.
For the rest of this proof, I'll abreviate the derivative of x^n as $(x^n)'$ and the derivative of x^{n-1} as $(x^{n-1})'$. Then using product rule, we know that:

$$(x^n)' = x(x^{n-1})' + 1 \cdot x^{n-1} = x \cdot (n-1)x^{n-2} + x^{n-1} = ((n-1)+1)x^{n-1} = nx^{n-1}$$

Claim 2: If f is differentiable at x and $\alpha \in \mathbb{R}$, then $(\alpha f)'(x) = \alpha f'(x)$.

Proof:

By the product rule: $(\alpha f)'(x) = \alpha f' + (\alpha)' f = \alpha f' + 0 \cdot f = \alpha f'$.

These combined with proposition 84 tells us that both polynomials and rational functions are differentiable over their domains.

Proposition 85: (chain rule)
Let f and g be real-valued functions with $\mathrm{dom}(f), \mathrm{dom}(g) \subseteq \mathbb{R}$. Let $x \in \mathbb{R}$. Suppose that f is differentiable at x and that g is differentiable at $f(x)$. Then $g \circ f$ is differentiable at x and $(g \circ f)'(x) = g'(f(x))f'(x)$.

Intuition:

$$\lim_{t \to x} \left(\frac{g(f(t)) - g(f(x))}{f(t) - f(x)} \cdot \frac{f(t) - f(x)}{t - x} \right) = g'(f(t)) \cdot f'(t).$$

That said, the issue with this intuition is that we need to address the possibility that $f(t) - f(x) = 0$.

Proof:

Set $y = f(x)$ and define $v(s) = \begin{cases} \frac{g(s)-g(y)}{s-y} - g'(y) & \text{if } s \neq y \\ 0 & \text{if } s = y \end{cases}$

Note that v is continuous at y. This is because g being differentiable at $f(x) = y$ means that:

$$\lim_{s \to y} v(s) = \lim_{s \to y} \left(\frac{g(s)-g(y)}{s-y} - g'(y) \right) = g'(y) - g'(y) = 0 = v(y).$$

Also, since f is differentiable at x, we know that f is continuous at x. Therefore, $v \circ f$ is continuous at x by proposition 68. Additionally, setting $s = f(t)$, we know that $s \to y$ as $t \to x$ because f is continuous at x. Thus:

$$\lim_{t \to x} v(f(t)) = \lim_{s \to y} v(s) = 0$$

Finally, note that $g(s) - g(y) = (s - y)(g'(y) + v(s))$ for all s. Thus by substituting that into our limit:

$$\begin{aligned}(g \circ f)'(x) &= \lim_{t \to x} \tfrac{g(f(t)) - g(f(x))}{t - x} \\ &= \lim_{t \to x} \tfrac{f(t) - f(x)}{t - x}(g'(f(x)) + v(f(t))) \\ &= f'(x)\,(g'(f(x)) + 0) \qquad \text{(by proposition 66)}\end{aligned}$$

Lecture 2: 4/3/2024

To start off lecture, here is some intuition about the behavior of derivatives. We'll formally define sine and cosine later (on page ___) but for this section please take for granted that $(\sin(x))' = \cos(x)$. Additionally, please take for granted that the power rule holds for non-positive integer exponents.

1. Define $f(x) = \begin{cases} x\sin(\frac{1}{x}) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$

   When $x \neq 0$, we have by chain rule that $f'(x) = \sin(\frac{1}{x}) - \frac{1}{x}\cos(\frac{1}{x})$.
   Meanwhile if $x = 0$, then $\frac{f(t) - f(0)}{t - 0} = \frac{t\sin(\frac{1}{t})}{t} = \sin(\frac{1}{t})$ when $t \neq 0$.
   So $\lim\limits_{t \to 0} \left(\frac{f(t) - f(0)}{t - 0}\right)$ does not exist, meaning f is not differentiable at x.

   This shows that $\mathrm{dom}(f')$ can be a proper subset of $\mathrm{dom}(f)$.

2. Define $g(x) = \begin{cases} x^2\sin(\frac{1}{x}) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$

   When $x \neq 0$, we have by chain rule that $g'(x) = 2x\sin(\frac{1}{x}) - \cos(\frac{1}{x})$.
   Meanwhile when $t \neq 0$:

   $$\left|\tfrac{g(t) - g(0)}{t - 0}\right| = \left|\tfrac{t^2\sin(\frac{1}{t})}{t}\right| = \left|t\sin(\tfrac{1}{t})\right| \leq |t|.$$

   Thus $0 = \lim\limits_{t \to 0}(-t) \leq \lim\limits_{t \to 0}\left(\frac{g(t) - g(0)}{t - 0}\right) \leq \lim\limits_{t \to 0}(t) = 0$, meaning $g'(0) = 0$.
   So $\mathrm{dom}(g') = \mathrm{dom}(g)$. That said, note that g' has a discontinuity of the second kind at 0. Therefore, this shows that the derivative of a function does not have to be continuous.

Let X be a metric space. A function $f : X \longrightarrow \mathbb{R}$ has a local maximum at $p \in X$ if $\exists \delta > 0 \;\; s.t. \;\; \forall x \in B_\delta(p), \;\; f(x) \leq f(p)$. Similarly, f has a local minimum if $\exists \delta > 0 \;\; s.t. \;\; \forall x \in B_\delta(p), \;\; f(x) \geq f(p)$.

Proposition 86: Let $f : (a,b) \longrightarrow \mathbb{R}$. If f has a local maximum at x and f is differentiable at x, then $f'(x) = 0$.

Proof:
Let $\delta > 0$ so that $\forall t \in B_\delta(x), \;\; f(t) \leq f(x)$. Then for all $t \in (x - \delta, x)$, $\frac{f(t)-f(x)}{t-x} \geq 0$. So $f'(x) \geq 0$. Similarly for all $t \in (x, x+\delta)$, we have $\frac{f(t)-f(x)}{t-x} \leq 0$. Thus $f'(x) \leq 0$.

Hence $f'(x) = 0$.

Note that analogous reasoning can show that if f has a local minimum at x and f is differentiable at x, then $f'(x) = 0$.

Proposition 87: If $f, g : [a,b] \longrightarrow \mathbb{R}$ are continuous on $[a,b]$ and differentiable on (a,b), then there exists $x \in (a,b)$ with:

$$(f(b) - f(a))g'(x) = (g(b) - g(a))f'(x).$$

Proof:
Define $h : [a,b] \longrightarrow \mathbb{R}$ by $h(x) = (f(b) - f(a))g(x) - (g(b) - g(a))f(x)$.
Then $h(a) = f(b)g(a) - g(b)f(a) = h(b)$.

Notice that h is continuous on $[a,b]$ and differentiable on (a,b) because of propositions 70 and 84. Since $h'(x) = (f(b)-f(a))g'(x)-(g(b)-g(a))f'(x)$, for all $x \in (a,b)$ it now suffices to show that there exists $x \in (a,b)$ with $h'(x) = 0$.

Since h is continuous on a compact set $[a,b]$, we know that h attains a maximum value and a minimum value over the interval $[a,b]$.

Case 1: If h is constant on $[a,b]$, then $h'(x) = 0$ for all $x \in (a,b)$.

Case 2: If there is $t \in (a,b)$ with $h(t) > h(a) = h(b)$, then $h(a)$ and $h(b)$ can't be the max. value that h attains on $[a,b]$. So h has a maximum at some point $x \in (a,b)$. Then by the last theorem, $h'(x) = 0$.

Case 3: If there is $t \in (a,b)$ with $h(t) < h(a) = h(b)$, then $h(a)$ and $h(b)$ can't be the min. value that h attains on $[a,b]$. So h has a minimum at some point $x \in (a,b)$. Then by the last theorem, $h'(x) = 0$.

Proposition 88: (Mean Value Theorem)
If $f : [a, b] \longrightarrow \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b), then there is $x \in (a, b)$ with $f(b) - f(a) = (b - a)f'(x)$.

To prove this, apply the previous proposition with $g(x) = x$.

Proposition 89: Suppose $f(a, b) \longrightarrow \mathbb{R}$ is differentiable. Then:

- If $f'(x) \geq 0$ for all $x \in (a, b)$, then f is monotone increasing.
- If $f'(x) \leq 0$ for all $x \in (a, b)$, then f is monotone decreasing.
- If $f'(x) = 0$ for all $x \in (a, b)$, then f is constant.

Proof:
For all $a < x_1 < x_2 < b$, we know by the mean value theorem that there exists $t \in (x_1, x_2)$ with $f(x_2) - f(x_1) = (x_2 - x_1)f'(t)$. Then since $x_2 - x_1 > 0$, the sign of $f(x_2) - f(x_1)$ depends entirely on $f'(t)$.

Lecture 3: 4/5/2024

Even though derivatives are not necessarily continuous, we can show they always satisfy the conclusion of the intermediate value theorem.

Proposition 90: Suppose $f : [a, b] \to \mathbb{R}$ is differentiable and $\lambda \in \mathbb{R}$ satisfies that $f'(a) < \lambda < f'(b)$. Then there is $x \in (a, b)$ with $f'(x) = \lambda$.

Proof:
Define $g : [a, b] \to \mathbb{R}$ by the rule $g(t) = f(t) - \lambda t$. Then g is differentiable with $g'(t) = f'(t) - \lambda$. So, it suffices to find $x \in (a, b)$ with $g'(x) = 0$

Since g is differentiable, we know that g is continuous. Adding in the fact that $[a, b]$ is compact, we know that g achieves a minimum value. So, let $x \in [a, b]$ be such that $g(x)$ is the minimum value of g.

Now consider that $f'(a) < \lambda < f'(b) \implies g'(a) < 0 < g'(b)$. Since $g'(a) < 0$, there is some $t_1 > a$ near a such that $g(x) \leq g(t_1) < g(a)$.

Explanation:
Set $\varepsilon = |g'(a)|$. Then by the definition of limits:

$$\exists \delta > 0 \ \ s.t. \ \ \forall t \in (a, a + \delta), \ \ \left| \frac{g(t)-g(a)}{t-a} - g'(a) \right| < \varepsilon.$$

Then because $g'(a)$ is negative, we must have that $\frac{g(t)-g(a)}{t-a} < 0$.
But as $t - a > 0$, we must have that $g(t) - g(a) < 0$.

This will be a common trick so get used to it.

Similarly, since $g'(b) > 0$, there is some $t_2 < b$ near b such that $g(x) \leq g(t_2) < g(b)$. Hence, we have shown that $x \neq a$ and $x \neq b$, meaning that $x \in (a,b)$. Then, by applying proposition 86 we know that $g'(x) = 0$.

We can prove an analogous theorem for when $f'(b) < \lambda < f'(a)$.

Corollary: If $f : [a,b] \longrightarrow \mathbb{R}$ is differentiable, then f' has no simple discontinuities.

Proof:
Assume that $x \in [a,b)$ and $f'(x+)$ exists. Then let $\varepsilon > 0$. By the definition of $f(x+)$:

$$\exists \delta > 0 \;\; s.t. \;\; \forall t \in (x, x+\delta), \;\; |f'(t) - f'(x+)| < \varepsilon/2.$$

If $f'(t) = f'(x)$ for all $t \in (x, x+\delta)$, then we automatically have that $f'(x+) = f'(x)$. So assume there exists $t \in (x, x+\delta)$ such that $f'(t) \neq f'(x)$. Then by the previous proposition, there exists $s \in (x,t)$ such that $f'(s)$ is between $f'(x)$ and $f'(t)$, and that $|f'(s) - f'(x)| < \varepsilon/2$.
Finally:

$$|f'(x) - f'(x+)| \leq |f'(x) - f'(s)| + |f'(s) - f'(x+)| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

So $f'(x)$ must equal $f'(x+)$. Similarly, we can show that if $x \in (a,b]$ and $f'(x-)$ exists, then $f'(x) = f'(x-)$. Thus, it is impossible for f' to have a simple discontinuity.

However, we already saw that f' can have discontinuities of the second kind.

Proposition 91: (L'Hôpital's rule)
Suppose $-\infty \leq a \leq b \leq +\infty$, that $f, g : (a,b) \longrightarrow \mathbb{R}$ are differentiable, and that $\forall x \in (a,b), \;\; g'(x) \neq 0$. Then suppose that $\lim\limits_{x \to a} \frac{f'(x)}{g'(x)} = A \in \mathbb{R} \cup \{-\infty, \infty\}$.
If either:

- both $f(x) \to 0$ and $g(x) \to 0$ as $x \to a$
- or either $g(x) \to +\infty$ or $g(x) \to -\infty$ as $x \to a$

then $\lim\limits_{x \to a} \frac{f(x)}{g(x)} \to A$.

(A similar result holds as $x \to b$.)

Proof:
Since $A \in \mathbb{R} \cup \{-\infty, \infty\}$, to show that $\lim\limits_{x \to a} \frac{f(x)}{g(x)} = A$, it suffices to show:

1. If $A \neq +\infty$, then for every $q \in \mathbb{R}$ with $q > A$, there is $c > a$ with $\forall x \in (a,c), \;\; \frac{f(x)}{g(x)} < q$.
2. If $A \neq -\infty$, then for every $q \in \mathbb{R}$ with $q < A$, there is $c > a$ with $\forall x \in (a,c), \;\; \frac{f(x)}{g(x)} > q$

Let's prove requirement 1. Assume $A \neq +\infty$ and fix $q \in \mathbb{R}$ with $q > A$. Next pick $r \in \mathbb{R}$ with $A < r < q$. Since $\frac{f'(x)}{g'(x)} \to A$ as $x \to a$, there is $c_1 > a$ with $\forall x \in (a, c_1),\ \ \frac{f'(x)}{g'(x)} < r$.

Now consider that whenever $a < x < y < c_1$, we have by proposition 87 that there exists $t \in (x, y)$ such that:

$$(f(y) - f(x))g'(t) = (g(y) - g(x))f'(t).$$

By the hypothesis of the theorem, $g'(t)$ can't be zero. Aditionally, because of the mean value theorem, if $g(y) - g(x) = 0$, then there would have to exist $s \in (x, y)$ with $g'(s) = 0$, thus contradicting the hypothesis of the theorem. So, it is safe to rearrange the above expression to get that:

$$\frac{f(y)-f(x)}{g(y)-g(x)} = \frac{f'(t)}{g'(t)} < r$$

Case 1: Assume $f(x) \to 0$ and $g(x) \to 0$ as $x \to a$.

Then fixing any $y \in (a, c_1)$, we have that $\lim\limits_{x\to a} \frac{f(y)-f(x)}{g(y)-g(x)} = \frac{f(y)}{g(y)} \leq r < q$.

Case 2: Assume $g(x) \to +\infty$ or $g(x) \to -\infty$ as $x \to a$.

Then fix any $y \in (a, c_1)$ and pick $c_2 \in (a, c_1)$ such that $\forall x \in (a, c_2)$, $g(x)$ and $g(x) - g(y)$ have the same sign. Then, $\forall x \in (a, c_2)$, we have that $\frac{g(x)-g(y)}{g(x)} > 0$. So:

$$\frac{f(y)-f(x)}{g(y)-g(x)} \cdot \frac{g(x)-g(y)}{g(x)} < r \cdot \frac{g(x)-g(y)}{g(x)}$$

Note that $\frac{f(y)-f(x)}{g(y)-g(x)} \cdot \frac{g(x)-g(y)}{g(x)} = \frac{f(x)-f(y)}{g(x)} = \frac{f(x)}{g(x)} - \frac{f(y)}{g(x)}$ and $\frac{g(x)-g(y)}{g(x)} = 1 - \frac{g(y)}{g(x)}$. Thus, we can rearrange terms to get that:

$$\frac{f(x)}{g(x)} < \left(1 - \frac{g(y)}{g(x)}\right) r + \frac{f(y)}{g(x)}$$

Now, $\lim\limits_{x\to a} \left(\left(1 - \frac{g(y)}{g(x)}\right) r + \frac{f(y)}{g(x)}\right) = (1-0)r + 0 = r$. So, there is $c_3 \in (a, c_2)$ such that $\forall x \in (a, c_3),\ \ \left(1 - \frac{g(y)}{g(x)}\right) r + \frac{f(y)}{g(x)} < q$.

Hence, $\forall x \in (a, c_3)$, $\frac{f(x)}{g(x)} < q$.

Requirement 2 is proved in a similar fashion. ■

Let f be a real-valued function with $\operatorname{dom}(f) \subseteq \mathbb{R}$. If f' is defined and is itself differentiable, then the derivative of f' is denoted f'' and called the <u>second derivative</u> of f. We similarly define $f''', f^{(4)}, \ldots, f^{(n)}$.

Also, we shall sometimes use $f^{(0)}$ to refer to the original function f.

Lecture 4: 4/8/2024

Proposition 92: (Taylor's Theorem)
Suppose that $f : [a, b] \longrightarrow \mathbb{R}$, that $f^{(n-1)}$ is continuous on $[a, b]$, and that $f^{(n)}$ is defined on (a, b). Then pick $\alpha \in [a, b]$ and define:

$$P(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t - \alpha)^k$$

Then for every $\beta \in [a, b] \setminus \{\alpha\}$, there is some x between α and β such that $f(\beta) = P(\beta) + \frac{f^{(n)}(x)}{n!}(\beta - \alpha)^n$.

Proof:
Set $M = \frac{f(\beta)-P(\beta)}{(\beta-\alpha)^n}$ so that $f(\beta) = P(\beta) + M(\beta - \alpha)^n$. Having done that, our goal is now to find an x between α and β such that $\frac{f^{(n)}(x)}{n!} = M$.

Define $g(t) = f(t) - P(t) - M(t - \alpha)^n$. Then, since P is a polynomial of degree $n - 1$, we have that $P^{(n)}(t) = 0$ for all t. So:

$$g^{(n)}(t) = f^{(n)}(t) - Mn!$$

Thus, it suffices to find an x between α and β such that $g^{(n)}(x) = 0$.

Importantly, P is the unique polynomial of degree $n - 1$ satisfying for all $0 \leq k \in\leq k - 1$ that $P^{(k)}(\alpha) = f^{(k)}(\alpha)$. Thus, for all $0 \leq k \in\leq k - 1$, we have that:

$$g^{(k)}(\alpha) = f^{(k)}(\alpha) - P^{(k)}(\alpha) - M\frac{n!}{(n-k)!}(\alpha - \alpha)^{n-k} = 0.$$

At the same time, for all $0 \leq k \leq n - 1$, we know that $g^{(k)}$ is continuous on $[\alpha, \beta]$ and differentiable on (α, β). So, we shall proceed by repeatedly applying the mean value theorem.

- $g(\beta) = 0$ and $g(\alpha) = 0$. So, there is x_1 between α and β with $g'(x_1) = 0$.
- $g'(x_1) = 0$ and $g'(\alpha) = 0$. So, there is x_2 between α and x_1 with $g''(x_2) = 0$.

$$\vdots$$

Eventually, you will get an x_n between α and x_{n-1} with $g^{(n)}(x_n) = 0$. ■

Note that this can be interpretted as a higher order analog of the mean value theorem. In fact, if $n = 1$ then this is just the mean value theorem.

The limit definition of the derivative still makes sense and can be applied to situations where f is a $\mathbb{C}$-valued or $\mathbb{R}^k$-valued function. Although, because this class is called "real" analysis, we shall always require that $\operatorname{dom}(f) \subseteq \mathbb{R}$.

(We will talk in 140C about when $\operatorname{dom}(f) \subseteq \mathbb{R}^k$)

If f is a $\mathbb{C}$-valued function, then we can write that $f = f_1 + if_2$ where f_1 and f_2 are real-valued. Then, f is differentiable if and only if f_1 and f_2 are differentiable. Also, $f'(x) = f_1'(x) + if_2'(x)$.

Proof:
Firstly consider any sequence (x_n) such that $x_n \to x$ as $n \to \infty$ but $x_n \neq x$ for any n. Then assuming $f'(x)$ exists, we know that:

$$\lim_{n\to\infty} \left| \frac{f(x_n)-f(x)}{x_n - x} - f'(x) \right| = 0$$

Now importantly:

- $0 \leq \left| \frac{f_1(x_n)-f_1(x)}{x_n-x} - \operatorname{Re}(f'(t)) \right| = \left| \operatorname{Re}\left(\frac{f(x_n)-f(x)}{x_n-x} - f'(x) \right) \right| \leq \left| \frac{f(x_n)-f(x)}{x_n-x} - f'(x) \right|$
- $0 \leq \left| \frac{f_2(x_n)-f_2(x)}{x_n-x} - \operatorname{Im}(f'(t)) \right| = \left| \operatorname{Im}\left(\frac{f(x_n)-f(x)}{x_n-x} - f'(x) \right) \right| \leq \left| \frac{f(x_n)-f(x)}{x_n-x} - f'(x) \right|$

So, $\lim\limits_{n\to\infty} \left| \frac{f_1(x_n)-f_1(x)}{x_n-x} - \operatorname{Re}(f'(x)) \right| = 0$ and $\lim\limits_{n\to\infty} \left| \frac{f_2(x_n)-f_2(x)}{x_n-x} - \operatorname{Im}(f'(x)) \right| = 0$.
This means $f_1'(x)$ and $f_2'(x)$ exist with $f_1'(x) = \operatorname{Re}(f'(x))$ and $f_2'(x) = \operatorname{Im}(f'(x))$.

Meanwhile, assume that $f_1'(x)$ and $f_2'(x)$ exist. Then:

$$\begin{aligned} f'(x) &= \lim_{t\to x} \left(\frac{f_1(t)+if_2(t)-f_1(x)-if_2(x)}{t-x} \right) \\ &= \lim_{t\to x} \left(\frac{f_1(t)-f_1(x)}{t-x} + i\frac{f_2(t)-f_2(x)}{t-x} \right) = f_1'(x) + if_2'(x). \end{aligned}$$

Similarly, if $\vec{f}$ is $\mathbb{R}^k$-valued, then we can write $\vec{f} = (f_1, f_2, \ldots, f_k)$ where $f_1, f_2, \ldots, f_k$ are real-valued. Then $\vec{f}$ is differentiable if and only if $f_1, f_2, \ldots, f_k$ are all differentiable. Also, $\vec{f}\,'(x) = (f_1'(x), f_2'(x), \ldots, f_k'(x))$.

This follows from the fact that given any sequence (x_n) such that $x_n \to x$ as $n \to \infty$ but $x_n \neq x$ for any n, we have by proposition 34 that: $\left(\frac{\vec{f}(x_n) - \vec{f}(x)}{x_n - x} \right)$ converges if and only if $\left(\frac{f_i(x_n)-f_i(x)}{x_n-x} \right)$ for each i.

For $\mathbb{C}$-valued functions, the addition, product, and quotient rules still hold.
For $\mathbb{R}^k$-valued functions, the addition and (dot) product rules still hold.

But, the mean value theorem and L'hôpital's rule fail in these situations.

For intuition on why this is, if f is $\mathbb{R}^k$ or $\mathbb{C}$-valued, then it is possible for $|f'|$ to be arbitrarily large over some interval of the domain while having f change as little as you want. To do this, make f "spin" in $\mathbb{R}^k$ or $\mathbb{C}$.

At least, we can still make the following theorem which is both similar to the mean value theorem and holds even for vector valued functions.

Proposition 93: Let $\vec{f} : [a, b] \longrightarrow \mathbb{R}^k$. Assume $\vec{f}$ is continuous on $[a, b]$ and differentiable on (a, b). Then there is $x \in (a, b)$ such that:

$$\|\vec{f}(b) - \vec{f}(a)\| \leq (b - a)\|\vec{f}'(x)\|$$

Proof:
Define $g : [a, b] \longrightarrow \mathbb{R}$ by $g(x) = (\vec{f}(b) - \vec{f}(a)) \cdot \vec{f}(x)$. Then g is continuous on $[a, b]$ and differentiable on (a, b). So by the mean value theorem there is $x \in (a, b)$ with $g(b) - g(a) = (b - a)g'(x)$.

Now note that:

$$\begin{aligned} |g(b) - g(a)| &= \left|\left(\vec{f}(b) - \vec{f}(a)\right) \cdot \vec{f}(b) - \left(\vec{f}(b) - \vec{f}(a)\right) \cdot \vec{f}(a)\right| \\ &= \left|\left(\vec{f}(b) - \vec{f}(a)\right) \cdot \left(\vec{f}(b) - \vec{f}(a)\right)\right| \\ &= \left\|\vec{f}(b) - \vec{f}(a)\right\|^2 \end{aligned}$$

Meanwhile, we also have that:

$$|g'(x)| = \left|\left(\vec{f}(b) - \vec{f}(a)\right) \cdot \vec{f}'(x)\right| \leq \left\|\vec{f}(b) - \vec{f}(a)\right\| \left\|\vec{f}'(x)\right\|$$

Therefore, we can combine equations to get that:

$$\begin{aligned} \|\vec{f}(b) - \vec{f}(a)\|^2 &= |g(b) - g(a)| \\ &= |b - a||g'(x) \leq |b - a|\|\vec{f}(b) - \vec{f}(a)\|\|f'(x)\| \end{aligned}$$

Now if $\vec{f}(b) - \vec{f}(a) = \vec{0}$, then this proposition is true trivially. So, it is safe to assume that $\|\vec{f}(b) - \vec{f}(a)\| \neq 0$. Then after canceling that, we get:

$$\|\vec{f}(b) - \vec{f}(a)\| \leq |b - a|\|f'(x)\|$$

Lecture 5: 4/10/2024

Now we move on to integrals…

To start, we define a partition of $[a, b]$ as a finite ordered set $P = \{x_0, x_1, \ldots, x_n\}$ with $a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$.

Note that in almost any other mathematical context, a partition means something else.

Here is how we define Riemann integrals:

Firstly, given a partition $P = \{x_0, x_1, \dots, x_n\}$, we write $\Delta x_i = x_i - x_{i-1}$ for each $i \in \{1, \dots, n\}$.

Now let $f : [a, b] \longrightarrow \mathbb{R}$ be a bounded function and $P = \{x_0, \dots, x_n\}$ be a partition of $[a, b]$. Then, we define for each $i \in \{1, \dots, n\}$:

- $m_i = \inf\{f(x) \mid x_{i-1} \leq x \leq x_i\}$
- $M_i = \sup\{f(x) \mid x_{i-1} \leq x \leq x_i\}$

Next, we define the lower estimate: $L(P, f) = \sum\limits_{i=1}^{n} m_i \Delta x_i$.

Simiarly, we define the upper estimate: $U(P, f) = \sum\limits_{i=1}^{n} M_i \Delta x_i$

Finally, letting $\mathcal{P}$ be the set of all partitions of $[a, b]$, we define:

($\mathcal{P}$ is not standard notation for that set. I just now made it up.)

- the lower Riemann integral as $\underline{\int_a^b} f dx = \sup\limits_{P \in \mathcal{P}} L(P, f)$
- the upper Riemann integral as $\overline{\int_a^b} f dx = \inf\limits_{P \in \mathcal{P}} U(P, f)$

And if $\underline{\int_a^b} f dx = \overline{\int_a^b} f dx$, then we denote the common value $\int_a^b f dx$ and call it the Riemann integral of f on $[a, b]$. Also, we call f Riemann integrable on $[a, b]$.

Some notes:

We write $\mathscr{R}_a^b$ to refer to the set of all functions that are Riemann integrable on $[a, b]$.

Also, since f is bounded, there are m and M with $\forall x \in [a, b]$, $m \leq f(x) \leq M$. Therefore, for every partition P:

$$m(b - a) \leq L(P, f) \leq U(P, f) \leq M(b - a).$$

So, $\underline{\int_a^b} f dx$ and $\overline{\int_a^b} f dx$ are always defined.

Meanwhile, here is how we define Riemann-Stieltjes integrals:

Let $\alpha : [a, b] \longrightarrow \mathbb{R}$ be monotone increasing. Then given a partition $P = \{x_0, x_1, \ldots, x_n\}$, we write $\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$ for each $i \in \{1, \ldots, n\}$.

Now let $f : [a, b] \longrightarrow \mathbb{R}$ be a bounded function and $P = \{x_0, \ldots, x_n\}$ be a partition of $[a, b]$. After defining m_i and M_i like before, we then define:

- the lower estimate: $L(P, f, \alpha) = \sum_{i=1}^{n} m_i \Delta\alpha_i$
- the upper estimate: $U(P, f, \alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_i$

Finally, we define:

- the lower Riemann-Stieltjes integral as $\underline{\int_a^b} f d\alpha = \sup_{P \in \mathcal{P}} L(P, f, \alpha)$
- the upper Riemann-Stieltjes integral as $\overline{\int_a^b} f d\alpha = \inf_{P \in \mathcal{P}} U(P, f, \alpha)$

And if $\underline{\int_a^b} f d\alpha = \overline{\int_a^b} f d\alpha$, then we denote the common value $\int_a^b f d\alpha$ and call it the Riemann-Stieltjes integral of f on $[a, b]$ with respect to α. Also, we call f Riemann-Stieltjes integrable on $[a, b]$ with respect to α.

Some notes:

We write $\mathscr{R}_a^b(\alpha)$ to refer to the set of all functions that are Riemann integrable on $[a, b]$ with respect to α.

Also, by defining $\alpha(x) = x$, we can see that the Riemann-Stieltjes integral is strictly more general than the Riemann integral.

Lecture 6: 4/12/2024

Let P_1 and P_2 be partitions of $[a, b]$. We say P_2 refines P_1 if $P_1 \subset P_2$.

Also, given two partitions P_1 and P_2 of $[a, b]$, their common refinement is $P_1 \cup P_2$ (reordered so that $x_i < x_{i+1}$ for all i).

Proposition 94: If P^* is a refinement of P, then for all bounded f and all monotone increasing α:

$$L(P,f,\alpha) \leq L(P^*,f,\alpha) \text{ and } U(P,f,\alpha) \geq U(P^*,f,\alpha).$$

Proof:
Firstly, since partitions are finite by definition, let's assume $P^* \setminus P$ consists of a single point x^*. After all, we can use induction to extend this result to when $|P^* \setminus P| > 1$. Also, let's focus on the lower estimate of P^* and P because the proof of this proposition for the upper estimate is mostly identical.

Say $P = \{x_0, x_1, \ldots, x_n\}$ and $x_{j-1} < x^* < x_j$ for some $j \in \{1, \ldots, n\}$. Then define $m_i = \inf\{f(x) \mid x_{i-1} \leq x \leq x_i\}$ for each $i \in \{1, \ldots, n\}$, $h_1 = \inf\{f(x) \mid x_{j-1} \leq x \leq x^*\}$, and $h_2 = \inf\{f(x) \mid x^* \leq x \leq x_j\}$.

Notice that $h_1, h_2 \geq m_j$. Also, we have that:

$$\alpha(x_j) - \alpha(x_{j-1}) = (\alpha(x_j) - \alpha(x^*)) + (\alpha(x^*) - \alpha(x_{j-1})).$$

Thus after canceling out many dupicate terms, we have that:

$$\begin{aligned} L(P^*,f,\alpha) - L(P,f,\alpha) &= h_2(\alpha(x_j) - \alpha(x^*)) + h_1(\alpha(x^*) - \alpha(x_{j-1})) - m_j(\alpha(x_j) - \alpha(x_{j-1})) \\ &= (h_1 - m_j)(\alpha(x^*) - \alpha(x_{j-1})) + (h_2 - m_j)(\alpha(x_j) - \alpha(x^*)) \\ &\geq 0 \text{ (because } \alpha \text{ is monotone increasing and } h_1, h_2 \geq m_j) \end{aligned}$$

So $L(P^*,f,\alpha) \geq L(P,f,\alpha)$.

Proposition 95: For every bounded f and monotone increasing α:

$$\underline{\int_a^b} f d\alpha \leq \overline{\int_a^b} f d\alpha$$

Proof:
If P_1 and P_2 are any partitions of $[a,b]$, then let P^* be their common refinement. Now obviously we have that $L(P^*,f,\alpha) \leq U(P^*,f,\alpha)$ because $m_i \leq M_i$ for each i. Combining that with the previous proposition, we have that:

$$L(P_1,f,\alpha) \leq L(P^*,f,\alpha) \leq U(P^*,f,\alpha) \leq U(P_2,f,\alpha)$$

Therefore, taking the supremum over all P_1, we get that:

$$\underline{\int_a^b} f d\alpha \leq U(P_2,f,\alpha) \text{ for all } P_2.$$

Then, taking the infimum over all P_2, we get that: $\underline{\int_a^b} f d\alpha \leq \overline{\int_a^b} f d\alpha$.

Proposition 96: Let f be bounded and α monotone increasing. Then:

$$f \in \mathscr{R}_a^b(\alpha) \iff \forall \varepsilon > 0, \ \ \exists P \ \ s.t. \ \ U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon.$$

Proof:
($\Longleftarrow$) For any partition P, we have that:

$$L(P, f, \alpha) \leq \underline{\int_a^b} f d\alpha \leq \overline{\int_a^b} f d\alpha \leq U(P, f, \alpha).$$

Thus, if $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$, we have that:

$$0 \leq \overline{\int_a^b} f d\alpha - \underline{\int_a^b} f d\alpha \leq U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon.$$

So, from the right-hand hypothesis we get that $\overline{\int_a^b} f d\alpha - \underline{\int_a^b} f d\alpha = 0$, which means that $f \in \mathscr{R}_a^b(\alpha)$.

($\Longrightarrow$) Assuming f is integrable gives us that $\underline{\int_a^b} f d\alpha = \int_a^b f d\alpha = \overline{\int_a^b} f d\alpha$.

So, let $\varepsilon > 0$ and pick two partitions P_1 and P_2 such that:

- $L(P_1, f, \alpha) > \underline{\int_a^b} f d\alpha - \varepsilon/2 = \int_a^b f d\alpha - \varepsilon/2$
- $U(P_2, f, \alpha) < \overline{\int_a^b} f d\alpha + \varepsilon/2 = \int_a^b f d\alpha + \varepsilon/2$

Then let P^* be the common refinement of P_1 and P_2. That way, when abbreviating $\int_a^b f d\alpha$ as the constant c, we have that:

$$c - \varepsilon/2 < L(P_1, f, \alpha) \leq L(P^*, f, \alpha) \leq U(P^*, f, \alpha) \leq U(P_2, f, \alpha) < c + \varepsilon/2.$$

So $U(P^*, f, \alpha) - L(P^*, f, \alpha) < \varepsilon$.

Proposition 97: Let f be bounded and α monotone increasing.

(A) If P^* refines P, then $U(P^*, f, \alpha) - L(P^*, f, \alpha) \leq U(P, f, \alpha) - L(P, f, \alpha)$.

Proof: This is just restating proposition 94.

Consider any partition $P = \{x_0, \ldots, x_n\}$ and for each $i \in \{1, \ldots, n\}$, pick $s_i, t_i \in [x_{i-1}, x_i]$.

(B) $\sum\limits_{i=1}^{n} |f(s_i) - f(t_i)| \, \Delta\alpha_i \leq U(P, f, \alpha) - L(P, f, \alpha)$

Proof:
For each i, we have that $f(s_i), f(t_i) \in [m_i, M_i]$. So:

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \, \Delta\alpha_i \leq \sum_{i=1}^{n} (M_i - m_i) \, \Delta\alpha_i = U(P, f, \alpha) - L(P, f, \alpha)$$

(C) If $f \in \mathscr{R}_a^b(\alpha)$, then $\left|\sum_{i=1}^{n} f(s_i)\Delta\alpha_i - \int_a^b f d\alpha\right| \leq U(P,f,\alpha) - L(P,f,\alpha)$.

Proof:
Since $m_i \leq f(s_i) \leq M_i$ for every i, we have that:

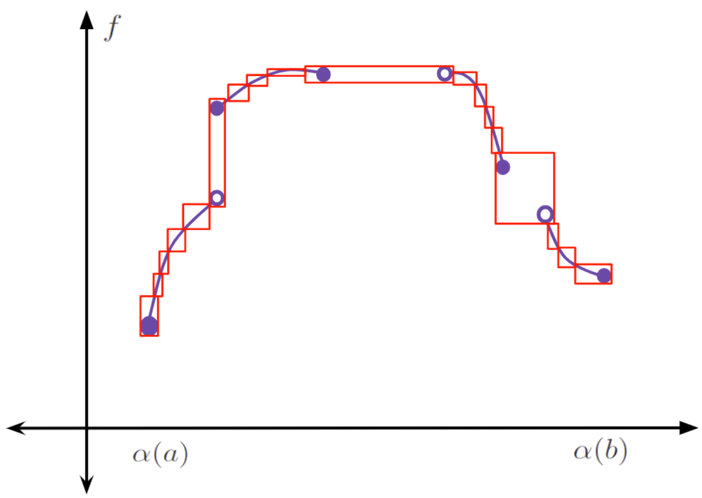
$$L(P,f,\alpha) \leq \sum_{i=1}^{n} f(s_i)\Delta\alpha_i \leq U(P,f,\alpha).$$

Also, we know that $L(P,f,\alpha) \leq \int_a^b f d\alpha \leq U(P,f,\alpha)$. Thus, combining these inequalities we get that:

$$\left|\sum_{i=1}^{n} f(s_i)\Delta\alpha_i - \int_a^b f d\alpha\right| \leq U(P,f,\alpha) - L(P,f,\alpha)$$

Lecture 7: 4/15/2024

Before, covering some sufficient conditions for integrability, it's worth going over some intuition for the next propositions.



Imagine the above parametric graph of $(\alpha(t), f(t))$ for $t \in [a,b]$. An important thing this diagram demonstrates is that both f and α can be discontinuous.

Given some partition, each area surrounded by a red rectangle corresponds to the quantity $(M_i - m_i)\Delta\alpha_i$. Also, the sum of all the areas in the red rectangles is $U(P,f,\alpha) - L(P,f,\alpha)$. So by proposition 96, we know that $f \in \mathscr{R}_a^b(\alpha)$ if and only if it possible to minimize the area inside those rectangles.

Observation:

- If α has a discontinuity at a point, then we are forced to have a wide rectangle.
- If f has a discontinuity at a point, then we are forced to have a tall rectangle.
- If both f and α are discontinuous at a point, then we're screwed because we're stuck having a wide and tall rectangle.

Proposition 98: If $f : [a, b] \longrightarrow \mathbb{R}$ is continuous and $\alpha : [a, b] \longrightarrow \mathbb{R}$ is monotonically increasing, then $f \in \mathscr{R}_a^b(\alpha)$.

Proof:

Recalling proposition 96, let $\varepsilon > 0$. Since f is continuous on the compact set $[a, b]$, it is uniformly continuous. So, there is $\delta > 0$ such that:

$$\forall x, t, \in [a, b], \quad |x - t| < \delta \Longrightarrow |f(x) - f(t)| < \tfrac{\varepsilon}{\alpha(b) - \alpha(a) + 1}$$

We know that $\alpha(b) - \alpha(a) \geq 0$. So, we add 1 to the denominator to make sure the denominator can not equal 0.

Now pick a partition $P = \{x_0, \ldots, x_n\}$ such that $\Delta x_i < \delta$ for all i. Then, for each i we have that $M_i - m_i \leq \frac{\varepsilon}{\alpha(b) - \alpha(a) + 1}$.

Technically, $M_i - m_i$ is strictly less than $\frac{\varepsilon}{\alpha(b) - \alpha(a) + 1}$ but we don't need that fact for this proof.

Then:

$$\begin{aligned} U(P, f, \alpha) - L(P, f, \alpha) &= \sum_{i=1}^{n} (M_i - m_i) \Delta\alpha_i \\ &\leq \tfrac{\varepsilon}{\alpha(b) - \alpha(a) + 1} \sum_{i=1}^{n} \Delta\alpha_i \\ &= \tfrac{\varepsilon}{\alpha(b) - \alpha(a) + 1} (\alpha(b) - \alpha(a)) \leq \varepsilon. \end{aligned}$$

Proposition 99: If $f : [a, b] \longrightarrow \mathbb{R}$ is monotone and $\alpha : [a, b] \longrightarrow \mathbb{R}$ is continuous and monotone increasing, then $f \in \mathscr{R}_a^b(\alpha)$.

Proof:

We'll assume that f is monotonically increasing because the proof for if f is monotonically decreasing is mostly identical.

Let $\varepsilon > 0$ and pick $n \in \mathbb{Z}_+$ big enough that $\frac{\alpha(b) - \alpha(a)}{n}(f(b) - f(a)) < \varepsilon$. Since α is continuous, by the intermediate value theorem we can find a partition $P = \{x_0, \ldots, x_n\}$ with $x_0 = a$, $x_n = b$, and

$$\alpha(x_i) = \alpha(a) + \tfrac{i}{n}(\alpha(b) - \alpha(a)).$$

Note then that $\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) = \frac{\alpha(b)-\alpha(a)}{n}$ for all i. Also, for all i we have that $m_i = f(x_{i-1})$ and $M_i = f(x_i)$ because f is monotonically increasing. Hence:

$$\begin{aligned} U(P,f,\alpha) - L(P,f,\alpha) &= \sum_{i=1}^{n}(M_i - m_i)\Delta\alpha_i \\ &= \tfrac{\alpha(b)-\alpha(a)}{n} \cdot \sum_{i=1}^{n}(f(x_i) - f(x_{i-1})) \\ &= \tfrac{\alpha(b)-\alpha(a)}{n}\,(f(b) - f(a)) < \varepsilon \end{aligned}$$

Proposition 100: Suppose $f : [a,b] \longrightarrow \mathbb{R}$ is bounded and has only finitely many discontinuities, and that $\alpha : [a,b] \longrightarrow \mathbb{R}$ is monotonically increasing and continuous at every point where f is discontinuous. Then $f \in \mathscr{R}_a^b(\alpha)$.

Proof:

Assume that $E = \{e_1, \ldots, e_m\}$ consists of all the points where f is discontinuous and define $M = \sup\{|f(x)| \mid a \leq x \leq b\}$. Let $\varepsilon > 0$. Since α is continuous at each $e_j \in E$, we can pick numbers $u_j < v_j$ for each $j \in \{1, \ldots, m\}$ such that:

1. $e_j \in [u_j, v_j]$
2. $e_j \in (u_j, v_j)$ when $e_j \notin \{a, b\}$
3. $\sum\limits_{j=1}^{m}(\alpha(v_j) - \alpha(u_j)) < \varepsilon$

Now set $K = [a,b] \setminus \bigcup\limits_{j=1}^{m}(u_j, v_j)$.

Importantly, K is a closed and bounded subset of $\mathbb{R}$, meaning it is compact. Also, f is continuous on K because K doesn't include any points where f is discontinuous except possibly a and b. But then, if f is discontinuous at a or b, then K includes a or b as an isolated point.

Therefore, f is uniformly continuous on K, meaning there exists $\delta > 0$ such that $\forall s,t \in K,\ \ |s-t| < \delta \Longrightarrow |f(s) - f(t)| < \varepsilon$. So, pick any partition $P = \{x_0, \ldots, x_n\}$ of $[a,b]$ such that:

- $\{u_1, v_1, u_2, v_2, \ldots, u_m, v_m\} \subseteq P \subseteq K$
- $x_{i-1} \notin \{u_1, u_2, \ldots, u_m\} \Longrightarrow \Delta x_i < \delta$ for all $i \in \{1, \ldots, n\}$

Also define $M_i, m_i, \Delta\alpha_i$ as usual using P and note that $M_i - m_i \leq 2M$. Additionally because $P \subseteq K$, we know that $[x_{i-1}, x_i] \subseteq K$ unless $x_{i-1} = u_j$ for some j. And if $x_{i-1} = u_j$, then $x_i = v_j$. Putting all of this together, we get that:

$$(M_i - m_i)\Delta\alpha_i \leq \begin{cases} 2M(\alpha(v_j) - \alpha(u_j)) & \text{if } x_{i-1} = u_j \text{ for some } j \\ \varepsilon\Delta\alpha_i & \text{if } x_{i-1} \notin \{u_1, \ldots, u_m\} \end{cases}$$

So:

$$\begin{aligned} U(P,f,\alpha) - L(P,f,\alpha) &= \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \\ &\leq 2M \sum_{j=1}^{m} (\alpha(v_j) - \alpha(u_j)) + \varepsilon(\alpha(b) - \alpha(a)) \\ &< (2M + \alpha(b) - \alpha(a))\varepsilon. \end{aligned}$$

$2M + \alpha(b) - \alpha(a)$ is a constant. Hence, it doesn't change the fact that we can make ε arbitrarily small and still find a partition P such that $U(P,f,\alpha) - L(P,f,\alpha) < \varepsilon$.

Lecture 8: 4/17/2024

Proposition 101: Suppose that $f : [a,b] \longrightarrow [m,M]$ such that $f \in \mathscr{R}_a^b(\alpha)$, and that $\phi : [m,M] \longrightarrow \mathbb{R}$ is continuous. Then $\phi \circ f \in \mathscr{R}_a^b(\alpha)$.

Proof:

Since ϕ is continuous and $[m,M]$ is compact, ϕ is uniformly continuous. So, there exists $\delta > 0$ such that $\delta < \varepsilon$ and:

$$\forall s,t \in [m,M], \;\; |s-t| < \delta \Longrightarrow |\phi(s) - \phi(t)| < \varepsilon.$$

Meanwhile, since $f \in \mathscr{R}_a^b(\alpha)$, we can pick a partition $P = \{x_0, \ldots, x_n\}$ of $[a,b]$ such that $U(P,f,\alpha) - L(P,f,\alpha) < \delta^2$. So, for each $i \in \{1,\ldots,n\}$, define m_i and M_i using P and f, as well as m_i^* and M_i^* using P and $\phi \circ f$. Then define $A = \{i \in \{1, \ldots n\} \mid M_i - m_i < \delta\}$ and $B = \{1,\ldots,n\} \setminus A$.

A is our "good" set and B is our "bad" set.

Now:

$$\begin{aligned} \delta \sum_{i \in B} \Delta\alpha_i \leq \sum_{i\in B} (M_i - m_i)\Delta\alpha_i &\leq \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \\ &= U(P,f,\alpha) - L(P,f,\alpha) < \delta^2. \end{aligned}$$

So, $\sum_{i\in B} \Delta\alpha_i < \delta$. At the same time, we have that $i \in A \Longrightarrow M_i^* - m_i^* \leq \varepsilon$.

Thus:

$$\begin{aligned} U(P,\phi\circ f,\alpha) - L(P,\phi\circ f,\alpha) &= \sum_{i=1}^{n} (M_i^* - m_i^*)\Delta\alpha_i \\ &= \sum_{i\in A} (M_i^* - m_i^*)\Delta\alpha_i + \sum_{i\in B} (M_i^* - m_i^*)\Delta\alpha_i \\ &\leq \varepsilon \sum_{i\in A} \Delta\alpha_i + 2\sup_{m\leq t\leq M} |\phi(t)| \cdot \sum_{i\in B} \Delta\alpha_i \\ &\leq \varepsilon(\alpha(b) - \alpha(a)) + 2\delta \sup_{m\leq t\leq M} |\phi(t)| \\ &\leq (\alpha(b) - \alpha(a) + 2\sup_{m\leq t\leq M} |\phi(t)|)\varepsilon \end{aligned}$$

Now $\alpha(b) - \alpha(a) + 2 \sup_{m \leq t \leq M} |\phi(t)|$ is a constant.

Thus, we've still shown that we can can make ε arbitrarily small and then find a partition P such that $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$.

Next, we explore the properties of integrals…

Proposition 102: Suppose that $c \in \mathbb{R}$, that $f, f_1, f_2 : [a, b] \longrightarrow \mathbb{R}$ are bounded, and that $\alpha, \alpha_1, \alpha_2 : [a, b] \longrightarrow \mathbb{R}$ be monotone increasing.

(A) $\underline{\int_a^b} f_1 d\alpha + \underline{\int_a^b} f_2 d\alpha \leq \underline{\int_a^b} (f_1 + f_2) d\alpha$ and $\overline{\int_a^b} f_1 d\alpha + \overline{\int_a^b} f_2 d\alpha \geq \overline{\int_a^b} (f_1 + f_2) d\alpha$

Hence, $f_1, f_2 \in \mathscr{R}_a^b(\alpha) \Longrightarrow (f_1 + f_2) \in \mathscr{R}_a^b(\alpha)$ with:

$$\int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha = \int_a^b (f_1 + f_2) d\alpha.$$

This follows from the fact that:

- $\inf_{x_{i-1} \leq x \leq x_i} f_1(x) + \inf_{x_{i-1} \leq x \leq x_i} f_2(x) \leq \inf_{x_{i-1} \leq x \leq x_i} (f_1(x) + f_2(x))$
- $\sup_{x_{i-1} \leq x \leq x_i} f_1(x) + \sup_{x_{i-1} \leq x \leq x_i} f_2(x) \geq \sup_{x_{i-1} \leq x \leq x_i} (f_1(x) + f_2(x))$

We proved this last quarter in a homework assignment. Intuitively though, this is because f_1 and f_2 might reach their most extreme values at different points, and in that case you can't maximize or minimize two functions at the same time.

Also, if $c \geq 0$, then $\underline{\int_a^b} cf d\alpha = c \underline{\int_a^b} f d\alpha$ and $\overline{\int_a^b} cf d\alpha = c \overline{\int_a^b} f d\alpha$. Meanwhile, if $c \leq 0$, then $\underline{\int_a^b} cf d\alpha = c \overline{\int_a^b} f d\alpha$ and $\overline{\int_a^b} cf d\alpha = c \underline{\int_a^b} f d\alpha$. Hence, we have that if $f \in \mathscr{R}_a^b(\alpha)$, then:

$$\int_a^b cf d\alpha = c \int_a^b f d\alpha.$$

This similarly follows from the fact that:

- when $c \geq 0$:
  - $\inf_{x_{i-1} \leq x \leq x_i} (cf(x)) = c \inf_{x_{i-1} \leq x \leq x_i} f(x)$
  - $\sup_{x_{i-1} \leq x \leq x_i} (cf(x)) = c \sup_{x_{i-1} \leq x \leq x_i} f(x)$
- when $c \leq 0$:
  - $\inf_{x_{i-1} \leq x \leq x_i} (cf(x)) = c \sup_{x_{i-1} \leq x \leq x_i} f(x)$
  - $\sup_{x_{i-1} \leq x \leq x_i} (cf(x)) = c \inf_{x_{i-1} \leq x \leq x_i} f(x)$

This is another past homework question.

(B) $\left(\forall x \in [a,b],\ \ f_1(x) \leq f_2(x)\right) \Longrightarrow \underline{\int_a^b} f_1 d\alpha \leq \underline{\int_a^b} f_2 d\alpha$ and $\overline{\int_a^b} f_1 d\alpha \leq \overline{\int_a^b} f_2 d\alpha$.

This is immediate from our definitions.

(C) If $c \in (a,b)$, then $\underline{\int_a^b} f d\alpha = \underline{\int_a^c} f d\alpha + \underline{\int_c^b} f d\alpha$ and $\overline{\int_a^b} f d\alpha = \overline{\int_a^c} f d\alpha + \overline{\int_c^b} f d\alpha$.

Proof:
We'll focus on the lower integral statement since the proof for the upper integral statement is mostly identical.

By proposition 94, we have that:

$$\underline{\int_a^b} f d\alpha = \sup_{P \in \mathcal{P}_{a,b}} L(P,f,\alpha) = \sup_{P \in \mathcal{P}_{a,b}} L(P \cup \{c\}, f, \alpha).$$

Next, setting $P_1 = (P \cup \{c\}) \cap [a,c]$ and $P_2 = (P \cup \{c\}) \cap [c,b]$, we have that:

$$\begin{aligned}\sup_{P \in \mathcal{P}_{a,b}} L(P \cup \{c\}, f, \alpha) &= \sup_{\substack{P_1 \in \mathcal{P}_{a,c} \\ P_2 \in \mathcal{P}_{c,b}}} \left(L(P_1, f, \alpha) + \sup_{P_2 \in \mathcal{P}_{c,b}} L(P_2, f, \alpha) \right) \\ &= \sup_{P_1 \in \mathcal{P}_{a,c}} L(P_1, f, \alpha) + \sup_{P_2 \in \mathcal{P}_{c,b}} L(P_2, f, \alpha) \\ &= \underline{\int_a^c} f d\alpha + \underline{\int_c^b} f d\alpha\end{aligned}$$

(D) $\left|\underline{\int_a^b} f d\alpha\right|, \left|\overline{\int_a^b} f d\alpha\right| \leq (\alpha(b) - \alpha(a)) \cdot \sup_{a \leq x \leq b} |f(x)|$.

This is immediate from our definitions.

(E) $\underline{\int_a^b} f d(\alpha_1 + \alpha_2) = \underline{\int_a^b} f d\alpha_1 + \underline{\int_a^b} f d\alpha_2$ and $\overline{\int_a^b} f d(\alpha_1 + \alpha_2) = \overline{\int_a^b} f d\alpha_1 + \overline{\int_a^b} f d\alpha_2$
Also, if $c \geq 0$, then $\underline{\int_a^b} c f d\alpha = c \underline{\int_a^b} f d\alpha$ and $\overline{\int_a^b} c f d\alpha = c \overline{\int_a^b} f d\alpha$.

This immediately follows from the fact that $\Delta(\alpha_1 + \alpha_2)_i = \Delta(\alpha_1)_i + \Delta(\alpha_2)_i$ and $\Delta(c\alpha)_i = c \cdot \Delta\alpha_i$.

Lecture 9: 4/19/2024

Proposition 103: If $f, g \in \mathscr{R}_a^b(\alpha)$, then:

(A) $fg \in \mathscr{R}_a^b(\alpha)$.

Proof:
By propositions 101 and 102, we know that $(f+g)^2 \in \mathscr{R}_a^b(\alpha)$ and $(f-g)^2 \in \mathscr{R}_a^b(\alpha)$. Therefore $fg = \frac{1}{4}\left((f+g)^2 + (f-g)^2\right) \in \mathscr{R}_a^b(\alpha)$.

(B) $|f| \in \mathscr{R}_a^b(\alpha)$ and $\left|\int_a^b f d\alpha\right| \le \int_a^b |f| d\alpha$.

Proof:
We know that $|f| \in \mathscr{R}_a^b(\alpha)$ by proposition 101. Then, because $f(x), -f(x) \le |f(x)|$ for all $x \in [a,b]$, we know by proposition 102.B that $\int_a^b f d\alpha \le \int_a^b |f| d\alpha$ and $-\int_a^b f d\alpha \le \int_a^b |f| d\alpha$. Therefore:
$$\left|\int_a^b f d\alpha\right| \le \int_a^b |f| d\alpha.$$

The unit step function I is defined by: $I(x) = \begin{cases} 0 & \text{when } x \le 0 \\ 1 & \text{when } x > 1 \end{cases}$

Proposition 104: Suppose $s \in [a,b)$ and $f : [a,b] \longrightarrow \mathbb{R}$ be bounded. If $\alpha(x) = I(x-s)$, we have that $f \in \mathscr{R}_a^b(\alpha)$ if and only if $f(s+) = f(s)$. Furthermore, $\int_a^b f d\alpha = f(s)$.

Proof:
Letting $P = \{x_0, \dots, x_n\}$ be any partition of $[a,b]$, we can assume because of proposition 94 that $s \in P$. So let j be such that $x_j = s$. Then note that $L(P,f,\alpha) = m_{j+1} = \inf\limits_{x_j \le x \le x_{j+1}} f(x)$ and $U(P,f,\alpha) = M_{j+1} = \sup\limits_{x_j \le x \le x_{j+1}} f(x)$.

($\Longrightarrow$) If $f \in \mathscr{R}_a^b$, then for all $\varepsilon > 0$, there exists a partition P such that $U(P,f,\alpha) - L(P,f,\alpha) < \varepsilon$. Then because $f(x), f(t) \in [m_{j+1}, M_{j+1}]$ for all $t \in [x_j, x_{j+1}]$, by setting $\delta = x_{j+1} - x_j$ we know that:
$$t \in [s, s+\delta] \Longrightarrow |f(t) - f(s)| \le M_{j+1} - m_{j+1} < \varepsilon.$$

So $f(s+) = f(s)$.

($\Longleftarrow$) Meanwhile if $f(s+) = f(s)$, then for any $\varepsilon > 0$ there exists $\delta > 0$ such that $t \in [s, s+\delta] \Longrightarrow |f(t) - f(s)| < \varepsilon/2$. So, let $P = \{x_0, \dots, x_n\}$ be a partition of $[a,b]$ such that $x_j = s$ and $x_{j+1} = s + \delta$. Then $m_{j+1} \ge f(s) - \varepsilon/2$ and $M_{j+1} \le f(s) + \varepsilon/2$. Hence:
$$U(P,f,\alpha) - L(P,f,\alpha) = M_{j+1} - m_{j+1} < \varepsilon$$

This means that $f \in \mathscr{R}_a^b(\alpha)$.

Finally, note that for all $\varepsilon > 0$ there exists a partition P such that $U(P,f,\alpha) = M_{j+1} < f(s) + \varepsilon$ and $L(P,f,\alpha) = m_{j+1} > f(s) + \varepsilon$.
Thus:
$$\underline{\int_a^b} f d\alpha = \overline{\int_a^b} f d\alpha = \int_a^b f d\alpha = f(s).$$

Proposition 105: Let (c_n) be a sequence of non-negative reals with $\sum c_n < \infty$. Also let (s_n) be a sequence in $[a,b)$ and set $\alpha(x) = \sum_{n=1}^{\infty} c_n I(x - s_n)$.

Interestingly, this is the same function we made at the bottom of page 63 during math 140A.

If $f : [a,b] \longrightarrow \mathbb{R}$ is continuous, then $f \in \mathscr{R}_a^b(\alpha)$ and $\int_a^b f d\alpha = \sum_{n=1}^{\infty} c_n f(s_n)$.

Proof:
Note that $\alpha(x)$ converges absolutely by comparison with (c_n) for all $x \in [a,b]$. Also, α is monotonically increasing with $\alpha(b) = \sum c_n$.

Now let $\varepsilon > 0$ and pick N with $\sum_{n=N+1}^{\infty} c_n < \varepsilon$. Also set $M = \sup_{a \leq x \leq b} |f(x)|$

By proposition 98, we know that $f \in \mathscr{R}_a^b(\alpha)$. So finally, define $\beta(x) = \sum_{i=1}^{N} c_k I(x - s_k)$ and $\gamma(x) = \sum_{i=N+1}^{\infty} c_k I(x - s_k)$. Then $\gamma = \alpha - \beta$ and $f \in \mathscr{R}_a^b(\beta), \mathscr{R}_a^b(\gamma)$. So:

$$\begin{aligned}
\left| \sum_{k=1}^{n} c_k f(s_k) - \int_a^b f d\alpha \right| &= \left| \int_a^b f d\beta - \int_a^b f d\alpha \right| \\
&= \left| \int_a^b f d(\beta - \alpha) \right| \\
&= \left| \int_a^b f d\gamma \right| \leq \int_a^b |f| d\gamma \\
&\leq \int_a^b M d\gamma = M(\gamma(b) - \gamma(a)) \\
&= M \left(\sum_{i=N+1}^{\infty} c_k - 0 \right) < M\varepsilon
\end{aligned}$$

As ε is arbitrary, we thus have that $\int_a^b f d\alpha = \sum_{n=1}^{\infty} c_n f(s_n)$.

Note, we assumed (c_n) and (s_n) are infinite sequences because if (c_n) and (s_n) have finitely many elements, then this theorem holds trivially because of proposition 102.

Proposition 106: Let $\alpha : [a,b] \longrightarrow \mathbb{R}$ be monotonically increasing and differentiable, and assume $\alpha' \in \mathscr{R}_a^b$. Also let $f : [a,b] \longrightarrow \mathbb{R}$ be bounded. Then $f \in \mathscr{R}_a^b(\alpha) \iff f\alpha' \in \mathscr{R}_a^b$ with $\int_a^b f d\alpha = \int_a^b f\alpha' dx$.

Proof:
It suffices to show $\underline{\int_a^b} f d\alpha = \underline{\int_a^b} f\alpha' dx$ and $\overline{\int_a^b} f d\alpha = \overline{\int_a^b} f\alpha' dx$.

We'll focus on proving the lower integral statement because the proof for the upper integral statement is mostly identical. Set $M = \sup_{a \le x \le b} |f(x)|$ and let $\varepsilon > 0$.

Pick partitions P_1, P_2, P_3 such that:

- $L(P_1, f, \alpha) > \underline{\int_a^b} f \, d\alpha - \varepsilon$
- $L(P_2, f\alpha') > \underline{\int_a^b} f\alpha' \, dx - \varepsilon$
- $U(P_3, \alpha') - L(P_3, \alpha') < \varepsilon$

Then set $P = P_1 \cup P_2 \cup P_3 = \{x_0, \ldots, x_n\}$ so that P satisfies all of the above inequalities.

By the mean value theorem, we know there exists $t_i \in [x_{i-1}, x_i]$ with $\Delta\alpha_i = \alpha'(t_i)\Delta x_i$ for each $1 \le i \le n$. Also, given any $s_i \in [x_{i-1}, x_i]$ for each $1 \le i \le n$, we know by proposition 97.B and because $U(P, \alpha') - L(P, \alpha') < \varepsilon$ that:

$$\sum_{i=1}^n |f(s_i) - f(t_i)|\Delta x_i < \varepsilon.$$

Therefore:

$$\begin{aligned}\left|\sum_{i=1}^n (f(s_i)\Delta\alpha_i - f(s_i)\alpha'(s_i)\Delta x_i)\right| &= \left|\sum_{i=1}^n (f(s_i)\alpha'(t_i)\Delta x_i - f(s_i)\alpha'(s_i)\Delta x_i)\right| \\ &\le M \cdot \left|\sum_{i=1}^n \alpha'(t_i)\Delta x_i - \sum_{i=1}^n \alpha'(s_i)\Delta x_i\right| < M\varepsilon\end{aligned}$$

Hence, letting $s_i = \inf_{x_{i-1} \le x \le x_i} f(x)$ for each $1 \le i \le n$,we have that:

$$\begin{aligned}\underline{\int_a^b} f \, d\alpha \ge L(P, f, \alpha) = \sum_{i=1}^n f(s_i)\Delta\alpha_i &> \sum_{i=1}^n f(s_i)\alpha'(s_i)\Delta x_i - M\varepsilon \\ &= L(P, f\alpha') - M\varepsilon > \overline{\int_a^b} f\alpha' \, dx - (M+1)\varepsilon\end{aligned}$$

Similarly, letting $s_i = \inf_{x_{i-1} \le x \le x_i} f(x)\alpha'(x)$ for each $1 \le i \le n$,we have that:

$$\begin{aligned}\underline{\int_a^b} f\alpha' \, dx \ge L(P, f\alpha') = \sum_{i=1}^n f(s_i)\alpha'(s_i)\Delta x_i &> \sum_{i=1}^n f(s_i)\Delta\alpha_i - M\varepsilon \\ &= L(P, f, \alpha) - M\varepsilon \\ &> \overline{\int_a^b} f\alpha' \, dx - (M+1)\varepsilon\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, we conclude that $\underline{\int_a^b} f \, d\alpha = \underline{\int_a^b} f\alpha' \, dx$.

Lecture 10: 4/22/2024

Proposition 107: Let $\alpha : [a,b] \longrightarrow \mathbb{R}$ be monotone increasing and let $f : [a,b] \longrightarrow \mathbb{R}$ with $f \in \mathscr{R}_a^b(\alpha)$. If $\phi : [A,B] \longrightarrow [a,b]$ is a strictly increasing continuous bijection, then $\int_A^B (f \circ \phi)\, d(\alpha \circ \phi) = \int_a^b f d\alpha$.

Proof:
Note that ϕ produces a bijection between partitions of $[A,B]$ and partitions of $[a,b]$. It sends the partition $Q = \{y_0, \ldots, y_n\}$ of $[A,B]$ to $P = \phi(Q) = \{x_0, \ldots, x_n\}$ where $x_i = \phi(y_i)$ for all $1 \le i \le n$.

With this notation, we have that:

$$\inf_{y_{j-1} \le t \le y_j} f \circ \phi(t) = \inf_{x_{j-1} \le t \le x_j} f(t) \text{ and } \sup_{y_{j-1} \le t \le y_j} f \circ \phi(t) = \sup_{x_{j-1} \le t \le x_j} f(t).$$

Also, $\Delta(\alpha \circ \phi)_i = \Delta\alpha_i$. So $U(Q, f \circ \phi, \alpha \circ \phi) = U(P, f, \alpha)$ and $L(Q, f \circ \phi, \alpha \circ \phi) = L(P, f, \alpha)$. Hence, the theorem above follows from taking the $\sup$ and $\inf$ over all Q.

Note: if $\alpha(x) = x$ and ϕ is differentiable with $\phi' \in \mathscr{R}_a^b$, then using the previous theorem we have that:

$$\int_a^b f(x)dx = \int_A^B (f \circ \phi) d\phi = \int_A^B f(\phi(x))\phi'(x)dx$$

Proposition 108: (part 1 of the fundamental theorem of calculus):
Let $f \in \mathscr{R}_a^b$ and define $F(x) = \int_a^x f(t)dt$ on the domain $[a,b]$. Then:

1. F is continuous.

   Proof:
   Let $\varepsilon > 0$ and set $M > 0$ such that $M \ge |f(x)|$ for all $x \in [a,b]$. Then set $\delta < \varepsilon/M$ and consider any $a \le x \le y \le b$ with $|x - y| < \delta$.

   By proposition 102, we know that:

   $$|F(y) - F(x)| = |\int_x^y f(t)dt| \le M|y - x| < M\delta < \varepsilon.$$

   So, F is uniformly continuous.

2. If f is continuous at x_0, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$.

   Proof:
   Let $\varepsilon > 0$ and pick $\delta > 0$ such that for all $t \in [a,b]$, $|t - x_0| < \delta$ implies that $|f(t) - f(x_0)| < \varepsilon$. Then consider any $s \ne x_0$ such that $s \in [a,b]$ and $|s - x_0| < \delta$.

   Notice that $f(x_0) = \frac{1}{|s - x_0|} \displaystyle\int_{\min(s,x_0)}^{\max(s,x_0)} f(x_0)dt$.

Therefore,

$$\left|\frac{F(s)-F(x_0)}{s-x_0} - f(x_0)\right| = \frac{1}{|s-x_0|}\left|\int_{\min(s,x_0)}^{\max(s,x_0)} (f(t)-f(x_0))\,dt\right|$$

$$\leq \frac{1}{|s-x_0|}\int_{\min(s,x_0)}^{\max(s,x_0)} |f(t)-f(x_0)|\,dt$$

$$< \frac{1}{|s-x_0|}\int_{\min(s,x_0)}^{\max(s,x_0)} \varepsilon dt = \frac{1}{|s-x_0|}|s-x_0|\varepsilon = \varepsilon$$

So $\lim\limits_{s\to x_0}\left|\frac{f(s)-F(x_0)}{s-x_0} - f(x_0)\right| = 0$, meaning that $F'(x_0) = f(x_0)$.

Proposition 109: (part 2 of the fundamental theorem of calculus)
If $f \in \mathscr{R}_a^b$ and there is a differentiable function $F : [a,b] \longrightarrow \mathbb{R}$ with $F' = f$, then $\int_a^b f dx = F(b) - F(a)$.

Proof:
Let $\varepsilon > 0$. Since $f \in \mathscr{R}_a^b$, there is a partition $P = \{x_0, \dots, x_n\}$ of $[a,b]$ with $U(P,f) - L(P,f) < \varepsilon$.

Now, by the mean value theorem, for each $1 \leq i \leq n$ pick $t_i \in [x_{i-1}, x_i]$ such that $F(x_i) - F(x_{i-1}) = F'(t_i)\Delta x_i = f(t_i)\Delta x_i$.

Then $F(b) - F(a) = \sum\limits_{i=1}^{n} f(t_i)\Delta x_i$ and $\sum\limits_{i=1}^{n} f(t_i)\Delta x_i$ is within ε of $\int_a^b f dx$ by proposition 97.C.

Proposition 110: (integration by parts)
If $F' = f \in \mathscr{R}_a^b$ and $G' = g \in \mathscr{R}_a^b$, then $Fg, fG \in \mathscr{R}_a^b$ with:

$$\int_a^b Fg dx = F(b)G(b) - F(a)G(a) - \int_a^b fG dx$$

Proof:
Since F and G are differentiable, they are continuous. Hence $F, G \in \mathscr{R}_a^b$. In turn, this means by proposition 103 that $Fg, fG \in \mathscr{R}_a^b$. So now, apply the fundamental theorem of calculus to $H(x) = F(x)G(x)$ and its derivative:

$$\int_a^b H' dx = \int_a^b (Fg + fG) dx = F(b)G(b) - F(a)G(a) = H(b) - H(a)$$

Lecture 11: 4/26/2024

If $\vec{f} : [a,b] \longrightarrow \mathbb{R}^k$ where $\vec{f} = (f_1, \dots, f_k)$ and $\alpha : [a,b] \longrightarrow \mathbb{R}$ is monotone increasing, then we say $\vec{f} \in \mathscr{R}_a^b(\alpha)$ if $\forall 1 \leq j \leq k, \;\; f_j \in \mathscr{R}_a^b(\alpha)$. Additionally, we define:

$$\int_a^b \vec{f}\, d\alpha = \left(\int_a^b f_1 d\alpha, \dots, \int_a^b f_k d\alpha \right)$$

Note that many of are previous theorems about integrals are immediately true about integrals about vector valued functions. This inclues proposition 102A, C, and E, as well as propositions 106, 108, and 109. To show this, just apply those propositions coordinatite-wise.

Proposition 111: If $\vec{f} \in \mathscr{R}_a^b(\alpha)$, then $\|\vec{f}\| \in \mathscr{R}_a^b(\alpha)$ and $\left\| \int_a^b f d\alpha \right\| \leq \int_a^b \|\vec{f}\| d\alpha$.

Proof:
Since $\vec{f} \in \mathscr{R}_a^b(\alpha)$ where $\vec{f} = (f_1, \dots, f_k)$, we know that $f_j \in \mathscr{R}_a^b(\alpha)$ for all $1 \leq j \leq k$. Hence, $\|\vec{f}\| = (f_1^2 + \dots + f_k^2)^{\frac{1}{2}} \in \mathscr{R}_a^b(\alpha)$ by propositions 101 and 102.

Now set $\vec{y} = \int_a^b \vec{f}\, d\alpha$ and say that $\vec{y} = (y_1, \dots, y_k)$. If $\vec{y} = \vec{0}$, there is nothing to prove. So assume $\vec{y} \neq \vec{0}$.

Next, by the Cauchy-Schwarz inequality we have that:

$$\sum_{j=1}^{k} y_j f_j(t) = \vec{y} \cdot \vec{f}(t) \leq \|\vec{y}\| \|\vec{f}(t)\|$$

So:

$$\begin{aligned} \|\vec{y}\| \left\| \int_a^b \vec{f}\, d\alpha \right\| = \left\| \int_a^b f d\alpha \right\|^2 & \\ = \vec{y} \cdot \int_a^b \vec{f}\, d\alpha = \sum_{j=1}^{k} y_j \int_a^b f_j d\alpha & \\ = \int_a^b \sum_{j=1}^{k} y_j f_j d\alpha \leq \int_a^b \|\vec{y}\| \|\vec{f}\| d\alpha & \\ = \|\vec{y}\| \int_a^b \|\vec{f}\| d\alpha & \end{aligned}$$

Thus, dividing by $\|\vec{y}\|$ we get that $\left\| \int_a^b f d\alpha \right\| \leq \int_a^b \|\vec{f}\| d\alpha$.

A curve is a continuous function $\gamma : [a,b] \longrightarrow \mathbb{R}^k$. We say γ is a closed curve if $\gamma(b) = \gamma(a)$. Also, we call γ a simple curve if γ is injective.

Given any partition $P = \{x_0, \ldots, x_n\}$ of $[a,b]$, we define:

$$\Lambda(P,\gamma) = \sum_{i=1}^{n} \|\gamma(x_i) - \gamma(x_{i-1})\|.$$

Then, the length of γ is $\Lambda(\gamma) = \sup\limits_{P \in \mathcal{P}} \Lambda(P,\gamma)$. When $\Lambda(\gamma)$ is finite, we say that γ is rectifiable.

Proposition 112: If γ is differentiable and γ' is continuous, then γ is rectifiable and $\Lambda(\gamma) = \int_a^b \|\gamma'(t)\| dt$.

Proof:
To start, note that $\gamma' \in \mathscr{R}_a^b$ since γ' is continuous. So by proposition 111, $\|\gamma'\| \in \mathscr{R}_a^b$. Now consider any partition $P = \{x_0, \ldots, x_n\}$ of $[a,b]$. By the fundamental theorem of calculation and proposition 111, we know:

$$\|\gamma(x_i) - \gamma(x_{i-1})\| = \left\| \int_{x_{i-1}}^{x_i} \gamma'(t) dt \right\| \leq \int_{x_{i-1}}^{x_i} \|\gamma'(t)\| dt$$

So for all partitions P:

$$\Lambda(P,\gamma) = \sum_{i=1}^{n} \|\gamma(x_i) - \gamma(x_{i-1})\| \leq \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} \|\gamma'(t)\| dt = \int_a^b \|\gamma'(t)\| dt$$

Therefore $\Lambda(\gamma) = \sup\limits_{P \in \mathcal{P}} \Lambda(P,\gamma) \leq \int_a^b \|\gamma'(t)\| dt$.

Meanwhile, let $\varepsilon > 0$. Since γ' is continuous on $[a,b]$, it is uniformly continuous. So:

$$\exists \delta > 0 \;\; s.t. \;\; \forall s,t \in [a,b], \;\; |s-t| < \delta \Longrightarrow \|\gamma'(s) - \gamma'(t)\| < \varepsilon.$$

Now pick a partition $P = \{x_0, \ldots, x_n\}$ with $\forall 1 \leq i \leq n, \;\; \Delta x_i < \delta$. Then:

$$\begin{aligned}
\int_{x_{i-1}}^{x_i} \|\gamma'(t)\| dt &\leq \int_{x_{i-1}}^{x_i} \|\gamma'(x_i)\| dt + \varepsilon \Delta x_i \\
&= \|\gamma'(x_i)\| \Delta x_i + \varepsilon \Delta x_i \\
&= \|\Delta x_i \gamma'(x_i)\| + \varepsilon \Delta x_i \\
&= \left\| \int_{x_{i-1}}^{x_i} \gamma'(x_i) dt \right\| + \varepsilon \Delta x_i \\
&= \left\| \int_{x_{i-1}}^{x_i} \gamma'(t) + \gamma'(x_i) - \gamma'(t) dt \right\| + \varepsilon \Delta x_i \\
&= \left\| \int_{x_{i-1}}^{x_i} \gamma'(t) dt \right\| + \varepsilon \Delta x_i + \varepsilon \Delta x_i \\
&= \|\gamma(x_i) - \gamma(x_{i-1})\| + 2\varepsilon \Delta x_i
\end{aligned}$$

So summing from $i = 1$ to n we get:

$$\int_a^b \|\gamma'(t)\| dt \leq \Lambda(P,\gamma) + 2\varepsilon(b-a) \leq \Lambda(\gamma) + 2\varepsilon(b-a).$$

And since $\varepsilon > 0$ was arbitary, we have that $\int_a^b \|\gamma'\| dt \leq \Lambda(\gamma)$.

Lecture 12: 4/29/2024

Suppose $f : E \longrightarrow \mathbb{C}$ where E is any set and (f_n) is a sequence of functions such that $f_n : E \longrightarrow \mathbb{C}$ for every $n \in \mathbb{N}$. We say that (f_n) converges to f pointwise on E if $\forall x \in E,\ \lim\limits_{n\to\infty} f_n(x) = f(x)$. Furthermore, in that case we call f the limit of (f_n).

Similarly, we write $f = \sum\limits_{n=1}^{\infty} f_n$ if $\forall x \in E,\ \ f(x) = \sum\limits_{n=1}^{\infty} f_n(x)$.

Questions:

1. If each f_n is continuous, is f continuous? Equivalently, we can ask:
$$\lim_{t\to x}\lim_{n\to\infty} f_n(t) = \lim_{t\to x} f(t) \stackrel{?}{=} f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty}\lim_{t\to x} f_n(t)$$

2. If each f_n is differentiable, is f differentiable and does f_n' converge to f'? Equivalently, we can ask:
$$\lim_{t\to x}\lim_{n\to\infty} \tfrac{f_n(t)-f_n(x)}{t-x} = \lim_{t\to x} \tfrac{f(t)-f(x)}{t-x} = f'(x) \stackrel{?}{=} \lim_{n\to\infty} f_n'(x) = \lim_{n\to\infty}\lim_{t\to x} \tfrac{f_n(t)-f_n(x)}{t-x}$$

3. If each $f_n \in \mathscr{R}_a^b$, is $f \in \mathscr{R}_a^b$ and does $\int_a^b f_n(x)dx$ converge to $\int_a^b f(x)dx$? Equivalently, we can ask:
$$\lim_{n\to\infty} \int_a^b f_n(x)dx \stackrel{?}{=} \int_a^b \left(\lim_{n\to\infty} f_n(x)\right) dx = \int_a^b f(x)dx$$

Some answers:

1. The order of limits generally matter:

   Set $s_{n,m} = \frac{n}{n+m}$. Then $\lim\limits_{m\to\infty}\left(\lim\limits_{n\to\infty} s_{n,m}\right) = \lim\limits_{m\to\infty}(1) = 1$ but $\lim\limits_{n\to\infty}\left(\lim\limits_{m\to\infty} s_{n,m}\right) = \lim\limits_{n\to\infty}(0) = 0$.

2. We can lose continuity:

   Set $f_n(x) = \frac{x^2}{(1+x^2)^n}$ and $f(x) = \sum\limits_{n=0}^{\infty} f_n(x)$. Then for all $n \in \mathbb{N}$, we have that $f_n(0) = 0$. So $f(0) = 0$.

   But when $x \neq 0$:
$$f(x) = \sum_{n=0}^{\infty} f_n(x) = x^2 \sum_{n=0}^{\infty} \tfrac{1}{(1+x^2)^n} = x^2 \sum_{n=0}^{\infty} \left(\tfrac{1}{1+x^2}\right)^n$$
$$= x^2 \frac{1}{1-\frac{1}{1+x^2}} = x^2 \cdot \tfrac{1+x^2}{x^2} = 1 + x^2$$

So despite every f_n being continuous, f is not continuous at $x = 0$.

3. We can lose differentiability:

   By the Stone-Weierstrass theorem (proposition 144), every continuous function is the limit of a sequence of polynomials. Therefore, as all polynomials are differentiable but not all continuous functions are differentiable, we know that there are sequences of differentiable functions whose limits are not differentiable.

4. We can lose integrability:

   Set $f_n(x) = \begin{cases} 1 & \text{if } x = \frac{p}{q} \text{ where } p, z \in Z \text{ and } |q| \leq n \\ 0 & \text{otherwise} \end{cases}$

   Then (f_n) converges pointwise to $f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{otherwise} \end{cases}$

   Also, because each f_n has finite discontinuities on $[-1, 1]$, we know $f_n \in \mathscr{R}_{-1}^{1}$. But $f \notin \mathscr{R}_{-1}^{1}$ by exercise 6.4.

5. Sometimes the derivatives of a sequence don't converge:

   Set $f_n(x) = \frac{\sin(nx)}{\sqrt{n}}$. Then (f_n) converges to $f(x) = 0$ and clearly $f'(x) = 0$. However, note that $f_n'(x) = \sqrt{n}\cos(nx)$. So (f_n') doesn't converge to $f'(x)$.

6. Sometimes the integrals of a sequence don't converge:

   Set $f_n(x) = nx(1-x^2)^n$ for $x \in [0, 1]$. Then clearly $f_n(0) = f_n(1) = 0$ for all $n \in \mathbb{N}$. Also, note that for any $0 < x < 1$, $\frac{1}{1-x^2} = 1 + \beta$ where $\beta > 0$. Thus by proposition 44.D (see 140A notes), we know that:

   $$n(1-x^2)^n = n^1 \frac{1}{\left(\frac{1}{1-x^2}\right)^n} = \frac{n^1}{(1+\beta)^n} \to 0 \text{ as } n \to \infty.$$

   So $f_n(x) \to 0$ for all $0 < x < 1$. This all shows that (f_n) converges pointwise to $f(x) = 0$, and clearly $\int_0^1 f(x)dx = 0$.

   Meanwhile, by the fundamental theorem of calculus we know that:

   $$\int_0^1 f_n(x)dx = \left[-\frac{n}{2(n+1)}(1-x^2)^{n+1}\right]_{x=0}^{x=1} = \frac{n}{2(n+1)}$$

   Therefore $\lim\limits_{n\to\infty} \int_a^b f_n(x)dx = \frac{1}{2} \neq 0 = \int_a^b f(x)dx$.

The take away of the previous examples is that pointwise convergence is annoying to work with. So here's a stronger notion of convergence that is more nice to work to with.

Once again suppose $f : E \longrightarrow \mathbb{C}$ where E is any set and (f_n) is a sequence of functions such that $f_n : E \longrightarrow \mathbb{C}$ for every $n \in \mathbb{N}$. We say that (f_n) converges to f uniformly on E if:

$$\forall \varepsilon > 0, \ \exists N \in \mathbb{N} \ \ s.t. \ \forall n \geq N \text{ and } \forall x \in E, \ \ |f_n(x) - f(x)| < \varepsilon$$

Similarly, $\sum\limits_{n=1}^{\infty} f_n$ converges uniformly to f on E if:

$$\forall \varepsilon > 0, \ \exists N \in \mathbb{N} \ \ s.t. \ \forall n \geq N \text{ and } \forall x \in E, \ \ |\sum_{k=1}^{n} f_k(x) - f(x)| < \varepsilon$$

Proposition 113: (Cauchy Criterion)
(f_n) converges uniformly on E if and only if:

$$\forall \varepsilon > 0, \ \exists N \in \mathbb{N} \ \ s.t. \ \forall n, m \geq N, \ \ \forall x \in E, \ \ |f_n(x) - f_m(x)| < \varepsilon.$$

Proof:
($\Longrightarrow$) Suppose (f_n) converges uniformly on E to $f : E \longrightarrow \mathbb{C}$ and let $\varepsilon > 0$. By definition there exists N such that for all $n \geq N$ and $x \in E$, we have that $|f_n(x) - f(x)| < \varepsilon/2$. Then whenever $n, m \geq N$, we have that:

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

($\Longleftarrow$) Suppose for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$ and $x \in E$, we have that $|f_n(x) - f_m(x)| < \varepsilon$. Then for every $x \in E$, the sequence $(f_n(x))$ is Cauchy, which means that it converges to some $f(x) \in \mathbb{C}$. In turn, this means that (f_n) converges to a function $f : E \longrightarrow \mathbb{C}$.

Now let $\varepsilon > 0$ and pick N satisfying that for all $n, m \geq N$ and $x \in E$, we have that $|f_n(x) - f_m(x)| < \varepsilon/2$. Then for any fixed $n \geq N$, we can take the limit as $m \to \infty$ in the above expression to get that:

$$\forall n \geq N, \ \ \forall x \in E, \ \ |f_n(x) - f(x)| \leq \varepsilon/2 < \varepsilon$$

Similarly, $\sum\limits_{n=1}^{\infty} f_n$ converges uniformly on E if and only if:

$$\forall \varepsilon > 0, \ \exists N \in \mathbb{N} \ \ s.t. \ \forall n, m \geq N, \ \ \forall x \in E, \ \ \left| \sum_{k=1}^{n} f_k(x) - \sum_{k=1}^{m} f_k(x) \right| < \varepsilon.$$

Proof:
This is just applying the above proposition to the sequence of functions:

$$\left(\sum_{k=1}^{n} f_k \right)_{n \in \mathbb{N}}$$

Proposition 114: Suppose (f_n) converges to $f : E \longrightarrow \mathbb{C}$. Then (f_n) converges uniformly on E to f if and only if $\lim_{n\to\infty} M_n = 0$ where $M_n = \sup_{x\in E} |f_n(x) - f(x)|$.

(This is just a restatement of our definition of uniform convergence.)

Lecture 13: 5/1/2024

Proposition 115: (Weierstrass M-test)

Let $f_n : E \longrightarrow \mathbb{C}$ be such that for all $n \in \mathbb{N}$, we have that $\forall x \in E$, $|f_n(x)| \leq M_n$. If $\sum_{n=1}^{\infty} M_n < \infty$, then $\sum_{n=1}^{\infty} f_n$ converges uniformly on E.

Proof:
Let $\varepsilon > 0$ and pick N such that $\forall m \geq n \geq N$, we have that $\sum_{i=n}^{m} M_i < \varepsilon$. Then for all $m \geq n \geq N$ and $x \in E$, we have that:

$$\left|\sum_{i=n}^{m} f_i(x)\right| \leq \sum_{i=n}^{m} |f_i(x)| \leq \sum_{i=n}^{m} M_i < \varepsilon$$

So by proposition 113, $\sum_{n=1}^{\infty} f_n$ converges uniformly on E.

Proposition 116: Suppose X is a metric space and (f_n) converges uniformly to f on $E \subseteq X$. Also let $x \in E'$ and assume for each n that $\lim_{t\to x} f_n(t) = A_n \in \mathbb{C}$. Then (A_n) converges with $\lim_{t\to x} f(t) = \lim_{n\to\infty} A_n$.

Proof:
Let $\varepsilon > 0$ and pick N such that for all $m \geq n \geq N$ and $t \in E$, we have that $|f_n(t) - f_m(t)| < \varepsilon/2$ (we can do this by proposition 113). Then for any $m \geq n \geq N$ we can take the limit as $t \to x$ to get that

$$|A_n - A_m| = \left|\lim_{t\to x} f_n(x) - \lim_{t\to x} f_m(x)\right| \leq \varepsilon/2 < \varepsilon$$

Therefore, (A_n) is Cauchy. And because $\mathbb{C}$ is complete, we thus know that $A_n \to A$ for some $A \in \mathbb{C}$.

Now let $\varepsilon > 0$. Since $A_n \to A$ and (f_n) converges uniformly on E, we can find N such that $|A_N - A| < \varepsilon/3$ and $\forall t \in E,\ \ |f_N(t) - f(t)| < \varepsilon/3$. At the same time, since $A_n = \lim_{t\to x} f_N(t)$, there exists $\delta > 0$ such that for all $t \in (B_\delta(x) \setminus \{x\}) \cap E,\ \ |f_N(t) - A_n| < \varepsilon/3$.

So, for all $t \in (B_\delta(x) \setminus \{x\}) \cap E$, we have that:

$$|f(t) - A| \leq |f(t) - f_N(t)| + |f_N(t) - A_n| + |A_n - A| < \varepsilon$$

Hence, $\lim_{t\to x} f(t) = A$.

Proposition 117: If X is a metric space, $E \subseteq X$, and $f_n : E \longrightarrow \mathbb{C}$ is continuous for each $n \in \mathbb{N}$, then if (f_n) converges uniformly to f on E we have that f is continuous.

Proof:

For any $x \in E$, if $x \notin E'$, then f is automatically continuous at x.

On the other hand, if $x \in E'$, then f is continuous at x if and only if $\lim\limits_{t \to x} f(t) = f(x)$. Luckily, by proposition 116 and because each f_n is continuous, we have that:

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

Proposition 118: Suppose that K is a compact metric space, that $f_n : K \longrightarrow \mathbb{R}$ are continuous, and that (f_n) converges to $f : K \longrightarrow \mathbb{R}$ pointwise. If f is continuous and $f_n(x) \leq f_{n+1}(x)$ for all $x \in K$ and $n \in \mathbb{N}$, then (f_n) converges uniformly to f on K.

Proof: Let $\varepsilon > 0$ and define $K_n = \{x \in K \mid f(x) - f_n(x) \geq \varepsilon\}$ for each $n \in \mathbb{N}$. Since f and f_n are continuous, we know that K_n is closed.

This is because $K_n = (f - f_n)^{-1}([\varepsilon, +\infty))$ and the preimage of a closed set with respect to a continuous function is always closed.

In turn, by proposition 24 (check 140A notes) we know that $K_n \subseteq K$ is compact for every $n \in \mathbb{N}$.

Now consider that because $f_n(x) \leq f_{n+1}(x)$ for all $x \in K$, we know that $K_{n+1} \subseteq K_n$. Additionally, since (f_n) converges to f pointwise, we know that for every $x \in K$ there exists $n \in \mathbb{N}$ such that $x \notin K_n$. This tells us that $\bigcap\limits_{n \in \mathbb{N}} K_n = \emptyset$.

So by the contrapositive of the finite intersection (check 140A notes), there is a finite collection of K_n such that $\bigcap\limits_{i=1}^{m} K_{n_i} = \emptyset$.

But, because $K_{n+1} \subseteq K_n$, for every n, this is only possible if one of the $K_{n_i} = \emptyset$. So, there exists N with $K_N = \emptyset$. Then for all $n \geq N$, $K_n = \emptyset$, which means that $\forall x \in K, \ |f(x) - f_n(x)| < \varepsilon$. So, (f_n) converges uniformly.

Let X be a metric space and $\mathscr{C}(X)$ be the the set of all continuous bounded functions from X to $\mathbb{C}$. Then for each $f \in \mathscr{C}(X)$, we define the uniform norm / supremum norm of f as:

$$\|f\| = \sup_{x \in X} |f(x)|$$

Importantly, $\|\cdot\|$ satisfies the properties of a norm:

- $\|f\|$ exists and is finite by the least upper bound property since $|f(x)|$ is bounded above and real.
- $\|f\|$ is non-negative because $|f(x)| \geq 0$ for all $x \in X$.
- $\|f\| = 0 \iff f(x) = 0$ for all $x \in X$.
- $\|f + g\| \leq \|f\| + \|g\|$ because $|f(x) + g(x)| \leq |f(x)| + |g(x)|$ for all $x \in X$.

In turn, this means that $\mathscr{C}(X)$ is a metric space when equipped with the metric:

$$d(f, g) = \|f - g\| = \sup_{x \in X} |f(x) - g(x)|$$

Also note that by proposition 114, a sequence (f_n) converges to f in this metric space if and only if (f_n) converges to f uniformly on X.

Proposition 119: $\mathscr{C}(X)$ is a complete metric space.

Proof:
If (f_n) is Cauchy in $\mathscr{C}(X)$, then by proposition 113, we know (f_n) converges uniformly to some function $f : X \longrightarrow \mathbb{C}$. Additionally, note that f is continuous by proposition 117. And because there is n with $\forall x \in X, \ \ |f_n(x) - f(x)| < 1$ and f_n is bounded, we know f is bounded as well. So, $f \in \mathscr{C}(X)$ and $\lim\limits_{n \to \infty} f_n = f$ in the metric space $\mathscr{C}(X)$.

Proposition 120: Let $\alpha : [a, b] \longrightarrow \mathbb{R}$ be monotone increasing and suppose that $f_n \to f$ uniformly on $[a, b]$ and $f_n \in \mathscr{R}_a^b(\alpha)$ for all $n \in \mathbb{N}$. Then $f \in \mathscr{R}_a^b \alpha$ and $\int_a^b f d\alpha = \lim\limits_{n \to \infty} \int_a^b f d\alpha$.

Proof:
Let $\varepsilon > 0$ and pick N such that for all $n \geq N$ and $x \in [a, b]$, we have that $|f_n(x) - f(x)| < \varepsilon$. Then fix any $n \geq N$.

We know that for all $x \in [a, b], \ \ f_n(x) - \varepsilon < f(x) < f_n(x) + \varepsilon$.
Therefore, by proposition 102 parts A and B, we have that:

$$\int_a^b f_n d\alpha - \varepsilon(\alpha(b) - \alpha(a)) \leq \underline{\int_a^b} f d\alpha \leq \overline{\int_a^b} f d\alpha \leq \int_a^b f_n d\alpha + \varepsilon(\alpha(b) - \alpha(a)).$$

Firstly, since $\varepsilon > 0$ is arbitrary, this shows that $f \in \mathscr{R}_a^b(\alpha)$. Secondly, we have that $\left| \int_a^b f_n d\alpha - \int_a^b f d\alpha \right| \leq \varepsilon(\alpha(b) - \alpha(a))$ for all $n \geq N$. Therefore, $\int_a^b f_n d\alpha \to \int_a^b f d\alpha$ as $n \to \infty$.

Corollary: If $f_n \in \mathscr{R}_a^b(\alpha)$ and $\sum_{n=1}^{\infty} f_n$ converges uniformly to f on $[a,b]$, then $\int_a^b f d\alpha = \sum_{n=1}^{\infty} \int_a^b f_n d\alpha$ (you can integrate term-by-term).

Lecture 14: 5/3/2024

Proposition 121: Assume $f_n : [a,b] \longrightarrow \mathbb{C}$ is differentiable for every $n \in \mathbb{N}$ and that there is a point $x_0 \in [a,b]$ such that $(f_n(x_0))$ converges. If (f_n') converges uniformly on $[a,b]$, then (f_n) converges uniformly on $[a,b]$ to some f satisfying that $f'(x) = \lim_{n\to\infty} f_n'(x)$ for all $x \in [a,b]$.

Proof:
We first want to check that (f_n) converges uniformly on $[a,b]$. So let $\varepsilon > 0$ and pick N so that $\forall n, m \geq N$, $|f_n(x_0) - f_m(x_0)| < \varepsilon/2$ and $|f_n'(t) - f_m'(t)| < \frac{\varepsilon}{2(b-a)}$ for all $t \in [a,b]$.

Then for any $n, m \geq N$ and any $x, y \in [a,b]$, we have by the mean value theorem that:

$$\begin{aligned}
&|f_n(x) - f_m(x) - f_n(y) + f_m(y)| \\
&\quad = |((f_n - f_m)(x) - (f_n - f_m)(y))| \\
&\quad\quad = |f_n'(t) - f_m'(t)||x-y| \quad \text{(for some } t \text{ between } x \text{ and } y\text{)} \\
&\quad\quad < \frac{\varepsilon}{2(b-a)}|x-y| = \frac{\varepsilon}{2} \cdot \frac{|x-y|}{b-a} < \frac{\varepsilon}{2} \cdot 1
\end{aligned}$$

Therefore, fixing y as x_0, we have that:

$$\begin{aligned}
&|f_n(x) - f_m(x)| \\
&\quad \leq |f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)| + |f_n(x_0) - f_m(x_0)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\end{aligned}$$

And so by proposition 113, we know that (f_n) converges uniformly.

Next, we want to show that $f'(x) = \lim_{n\to\infty} f_n'(x)$ for all $x \in [a,b]$. So, fix any $x \in [a,b]$ and define:

$$\phi(t) = \frac{f(t)-f(x)}{t-x} \text{ and } \phi_n(t) = \frac{f_n(t)-f_n(x)}{t-x} \text{ for all } t \in [a,b] \setminus \{x\}.$$

Clearly (ϕ_n) converges to ϕ pointwise. Also, note that for any $n, m \geq N$:

$$|\phi_n(t) - \phi_m(t)| = \frac{|f_n(t)-f_n(x)-f_m(t)+f_m(x)|}{|t-x|} < \frac{\varepsilon}{2} \cdot \frac{|t-x|}{b-a} \cdot \frac{1}{|t-x|} = \frac{\varepsilon}{2(b-a)}$$

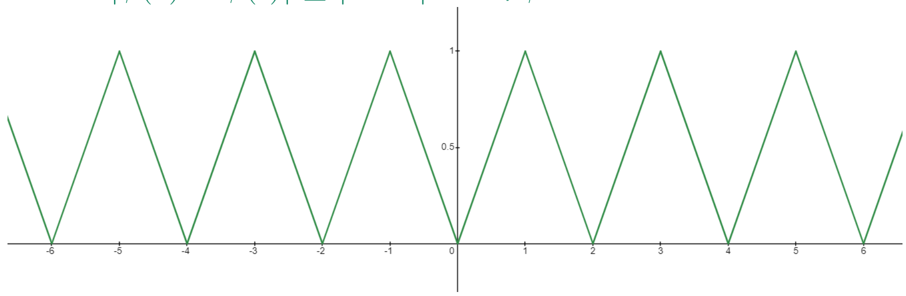
Thus we also know by proposition 113 that (ϕ_n) converges uniformly to ϕ on $[a,b] \setminus \{x\}$. And so by proposition 116, we know that:

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

Proposition 122: There exists a function $f : \mathbb{R} \longrightarrow \mathbb{R}$ that is continuous but nowhere differentiable.

Proof:
Define $f(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \phi(4^n x)$ where $\phi(x)$ is the distance from x to the nearest even integer.

As is hopefully visably apparent from the graph below, for all $s, t \in \mathbb{R}$ we have that $|\phi(s) - \phi(t)| \leq |s - t|$. Also, ϕ is continuous.



Now note that $\sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \phi(4^n x)$ converges uniformly by the Weierstrass M-test.

Set $M_n = \left(\frac{3}{4}\right)^n$. Then $\sum_{n=0}^{\infty} M_n = \frac{1}{1-\frac{3}{4}} = 4$ and $M_n \geq \left|\left(\frac{3}{4}\right)^n \phi(4^n x)\right|$ for all $x \in \mathbb{R}$.

Thus, since $\sum_{k=1}^{n} \left(\frac{3}{4}\right)^n \phi(4^n x)$ is continuous for each $n \in \mathbb{N}$, we have by proposition 117 that f is continuous on $\mathbb{R}$.

Next, fix $x \in \mathbb{R}$ and $m \in \mathbb{N}$. Then set $\delta_m = \pm\frac{1}{2} \cdot 4^{-m}$ (choosing the sign so that no integer is stirctly between $4^m x$ and $4^m(x + \delta_m)$). For any $n \in \mathbb{N}$, set $\gamma_n = \frac{\phi(4^n(x+\delta_m)) - \phi(4^n x)}{\delta_m}$.

Then $|\gamma_n| \begin{cases} \leq 4^m & \text{if } n < m \\ = 4^m & \text{if } n = m \\ = 0 & \text{if } n > m \end{cases}$

So we have that $\left|\frac{f(x+\delta_m) - f(x)}{\delta_m}\right| = \left|\sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \gamma_n\right| = \left|\sum_{n=0}^{m} \left(\frac{3}{4}\right)^n \gamma_n\right|$.

Now since $\gamma_m = 4^m$, we have in the absolute worst case that:

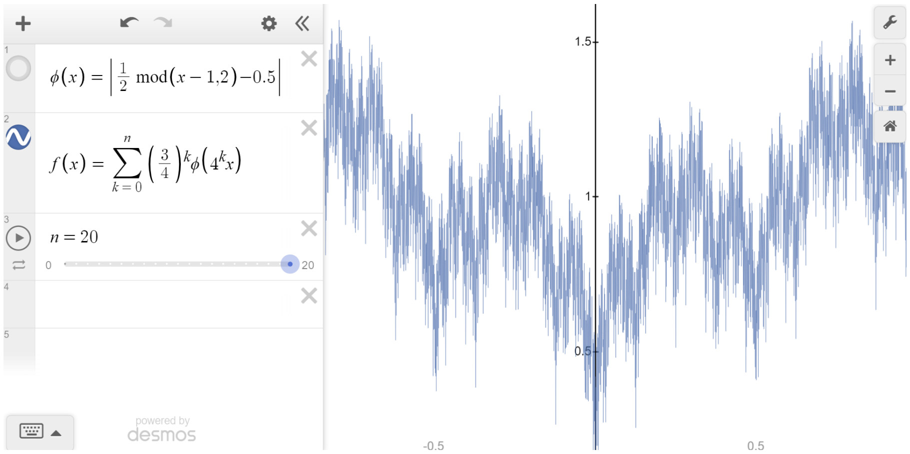
$$\left|\sum_{n=0}^{m} \left(\frac{3}{4}\right)^n \gamma_n\right| \geq 3^n - \sum_{n=0}^{m-1} \left|\left(\frac{3}{4}\right)^n \gamma_n\right| \geq 3^n - \sum_{n=0}^{m-1} 3^n$$

$$= 3^m - \frac{3^m - 1}{3 - 1} = 3^m - \frac{1}{2}\left(3^m - 1\right) = \frac{1}{2}3^m + \frac{1}{2}$$

As $m \to \infty$, $\delta_m \to 0$. So this shows that $\lim_{t \to x} \frac{f(t) - f(x)}{t - x}$ does not exist in $\mathbb{R}$ since $\frac{1}{2}3^m + \frac{1}{2} \to \infty$ as $m \to \infty$.

Since x was arbitrary, we thus know that f is nowhere differentiable.

Note:
Here is what this function looks like in Desmos:



Obviously, I can't graph an infinite sum. But hopefully the first 20 terms are enough to give you an idea of what this function looks like.

Lecture 15: 5/7/2024

We know that every bounded sequence in $\mathbb{R}$ has a convergent subsequence. But is something similar true of functions?

For all $n \in \mathbb{N}$, let $f_n : E \longrightarrow \mathbb{C}$.

- We say (f_n) is pointwise bounded on E if for every fixed $x \in E$, the sequence $(f_n(x))$ is bounded. Equivalently, we can say that f is pointwise bounded if there is $\phi : E \longrightarrow \mathbb{R}$ such that for all $x \in E$ and $n \in \mathbb{N}$, $|f_n(x)| \leq \phi(x)$.
- We say (f_n) is uniformly bounded on E if there is $M \geq \mathbb{R}$ with for all $x \in E$ and for all $n \in \mathbb{N}$, $|f_n(x)| \leq M$.

Now in really restricted situations, (f_n) is guarenteed to have a convergent subsequence.

Proposition 123: If E is countable, $f_n : E \longrightarrow \mathbb{C}$ for all $n \in \mathbb{N}$, and (f_n) is pointwise bounded, then there is a subsequence (f_{n_k}) that converges pointwise on E.

Proof:
Since E is countable, we can enumerate E as $E = \{x_1, x_2, \ldots\}$. Then, set $I_0 = \mathbb{N}$ and inductively assume that $I_{k-1} \subseteq \mathbb{N}$ has been defined and is infinite.

Since $(f_n(x_k))_{n\in I_{k-1}}$ is bounded, there is a convergent subsequence. Hence, there is an infinite set $I_k \subseteq I_{k-1}$ so that $(f_n(x_k))_{n\in I_k}$ converges. This defines I_k for all $k \in \mathbb{N}$.

Now, define n_k to be the kth element of I_k. Importantly, this is a well defined subsequence of $\mathbb{N}$ because $n_k > n_{k-1}$ for every $k \in \mathbb{N}$ (we know this because $I_k \subseteq I_{k-1}$).

We claim $(f_{n_k})_{k\in\mathbb{N}}$ converges pointwise on E. Indeed, for any $m \geq 1$, we have that $\forall k \geq m,\ \ n_k \in I_k \subseteq I_m$. So, $(f_{n_k}(x_m))_{k\geq M}$ is a subsequence of the convergent sequence $(f_n(x_m))_{n\in I_m}$ and thus $(f_{n_k}(x_m))$ converges for every $x_m \in E$.

Often however, (f_n) has no convergent subsequence.

Here is a uniformly bounded sequence with no pointwise convergent subsequence:

Set $f_n(x) = \sin(nx)$ and suppose towards a contradiction there is a subsequence (n_k) with $(f_{n_k})_{k\in\mathbb{N}}$ pointwise convergent on $[0, 2\pi]$. Then for all $x \in [0, 2\pi]$, $\ (\sin(n_k x) - \sin(n_{k+1}x))^2 \to 0$. So by theorem 11.32 (check 140C notes), $\int_0^{2\pi}(\sin(n_k x) - \sin(n_{k+1}x))^2 dx \to 0$.

However, $\int_0^{2\pi}(\sin(n_k x) - \sin(n_{k+1}x))^2 dx = 2\pi$. Thus we have a contradiction.

Here is a uniformly bounded pointwise convergence sequence on a compact set with no uniformly convergent subsequence:

Set $f_n(x) = \frac{x^2}{x^2+(1-nx)^2}$ for $x \in [0, 1]$.

Then $\forall x, \forall n,\ \ |f_n(x)| \leq 1$. So (f_n) is uniformly bounded.

Also, $\forall x,\ \ \lim\limits_{n\to\infty} f_n(x) = 0$. So (f_n) is pointwise convergent.

However, $f_n(\frac{1}{n}) = 1$ for all $n \in \mathbb{N}$. So, there can't be a uniformly convergent subsequence.

A collection $\mathcal{F}$ of functions $f : E \longrightarrow \mathbb{C}$ is equicontinuous if

$$\forall \varepsilon > 0,\ \ \exists \delta > 0\ \ s.t.\ \ \forall f \in \mathcal{F},\ \ \forall x, y \in E,\ \ d(x, y) < \delta \Longrightarrow |f(x) - f(y)| < \varepsilon$$

Proposition 124: If K is a compact metric space and $f_n \in \mathscr{C}(K)$ for each $n \in \{1, 2, \ldots,\}$ such that (f_n) converges uniformly on K, then (f_n) is equicontinuous on K.

Proof:

Let $\varepsilon > 0$. Then since (f_n) converges uniformly, there is an integer N such that $\|f_n - f_N\| < \varepsilon/3$ for all $n > N$.

(This is by the Cauchy Criterion)

Also, since each f_n is uniformly continuous on K, there exists $\delta > 0$ such that $|f_i(x) - f_i(y)| < \varepsilon/3$ for each $1 \leq i \leq N$ and $d(x,y) < \delta$.

In turn, for all $n > N$ and $d(x,y) < \delta$, we also have that

$$|f_n(x) - f_n(y)| \\ \leq |f_n(x) - F_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| < \varepsilon$$

And thus (f_n) is equicontinuous.

Lecture 16: 5/8/2024

Proposition 125: Suppose K is compact, $f_n \in \mathscr{C}(K)$ for each $n \in \mathbb{N}$, and (f_n) is pointwise bounded and equicontinuous on K. Then:

(A) (f_n) is uniformly bounded.

Proof:
Pick a $\delta > 0$ using equicontinuity so that for all $n \in N$ when $d(x,y) < \delta$, $\;|f_n(x) - f_n(y)| < 1$. Then, since K is compact, we can find $p_1, \ldots, p_m \in K$ with $\bigcup\limits_{i=1}^{m} B_\delta(p_i) \supseteq K$.

Also, since (f_n) is pointwise bounded, $\forall 1 \leq i \leq m$, we know that $(f_n(p_i))$ is bounded. So there is $M > 0$ with $\forall 1 \leq i \leq m$, $\forall n \in \mathbb{N},\;\; |f_n(p_i)| \leq M$.

Now consider any $x \in K$ and pick $i \in \{1, \ldots, m\}$ with $x \in B_\delta(p_i)$. Then for any n we have:

$$|f_n(x)| \leq |f_n(p_i)| + |f_n(p_i) - f_n(k)| < M + 1$$

Therefore, (f_n) is uniformly bounded.

(B) There is a subsequence (f_{n_i}) that converges uniformly on K.

Proof:
To start, we proved in a homework question for 140A that if K is compact, then there exists a countable set E that is dense in K. Also, by proposition 123, there is a subsequence (f_{n_i}) that converges pointwise on E. Now, setting $g_i = f_{n_i}$ we will show that (g_i) converges uniformly on K.

Let $\varepsilon > 0$ and pick $\delta > 0$ using equicontinuity such that:

$$\forall i \in \mathbb{N},\;\; \forall x,t \in K,\;\; d(x,t) < \delta \implies (g_i(x) - g_i(t)) < \varepsilon/3$$

Since E is dense in K and K is compact, there are $p_1, \ldots, p_m \in E$ with $\bigcup_{l=1}^{m} B_\delta(P_l) \supseteq K$.

Also, since $\forall 1 \leq l \leq m$, we have that $(g_i(p_l))$ converges, there is N such that $\forall i, j \geq N, \ \ \forall 1 \leq l \leq m, \ \ |g_i(p_l) - g_j(p_l)| < \varepsilon/3$. And, for any $x \in K$ and $i, j \geq N$, we can pick $1 \leq l \leq m$ with $x \in B_\delta(p_l)$. Then:

$$|g_i(x) - g_j(x)| \leq |g_i(x) - g_i(p_l)| + |g_i(p_l) - g_j(p_l)| + |g_j(p_l) - g_j(x)| < \varepsilon$$

Lecture 17: 5/10/2024

Proposition 126: (Weierstrass Theorem)
If $f : [a, b] \longrightarrow \mathbb{C}$ is continuous, then there is a sequence (P_n) of polynomials such that (P_n) converges to f uniformly on $[a, b]$. Also, if f is $\mathbb{R}$-valued, each P_n can be chosen to be $\mathbb{R}$-valued.

Proof:
To start, without loss of generality we can assume that $[a, b] = [0, 1]$ and $f(0) = f(1) = 0$.

This is because given any $f : [a, b] \longrightarrow \mathbb{C}$, we can define $g : [0, 1] \longrightarrow \mathbb{C}$ by $g(x) = f(x(b - a) + a) - f(a) - x(f(b) - f(a))$.

Then, if there exists (P_n) converging uniformly to g on $[0, 1]$, we know that $P_n(\frac{x-a}{b-a}) + f(a) + \frac{x-a}{b-a}(f(b) - f(a))$ converges uniformly on $[a, b]$ to $g(\frac{x-a}{b-a}) + f(a) + \frac{x-a}{b-a}(f(b) - f(a)) = f(x)$.

So, extend f to all of $\mathbb{R}$ by setting $f(x) = 0$ when $x \in \mathbb{R} \setminus [0, 1]$. Note that f is the uniformly continuous on $\mathbb{R}$.

Claim 1: There exist polynomials $Q_n : [-1, 1] \longrightarrow [0, \infty)$ such that

(A) $\int_{-1}^{1} Q_n(t)dt = 1$

(B) $\forall \delta \in (0, 1), \ \ \lim_{n \to \infty} \left(\int_{-1}^{-\delta} Q_n(t)dt + \int_{\delta}^{1} Q_n(t)dt \right) = 0$

Proof:
Set $Q_n(x) = c_n(1 - x^2)^n$ where $c_n = \left(\int_{-1}^{1} (1 - t^2)^n dt \right)^{-1}$. Then each Q_n satisfies claim 1.A.

Also, for each $x \in [-1, 1], \ \ Q_n(x) \geq 0$.

Additionally, since $(1 - x^2)^n - (1 - nx^2)$ is even, zero at $x = 0$, and has a positive derivative on $(0, 1)$, we have that $\forall x \in [-1, 1]$, $(1 - x^2)^n \geq 1 - nx^2$.

So:

$$\begin{aligned}\int_{-1}^{1}(1-x^2)^n dx &= 2\int_0^1 (1-x^2)^n dx\\ &\geq 2\int_0^{1/\sqrt{n}}(1-x^2)^n dx\\ &\geq 2\int_0^{1/\sqrt{n}}(1-nx^2)dx = 2\left(\frac{1}{\sqrt{n}}-\frac{n}{3}\left(\frac{1}{\sqrt{n}}\right)^3\right)\\ &= 2\left(\frac{1}{\sqrt{n}}-\frac{1}{3\sqrt{n}}\right) = \frac{1}{\sqrt{n}}\cdot\frac{4}{3} > \frac{1}{\sqrt{n}}\end{aligned}$$

This tells us that each $c_n < \sqrt{n}$.

Now finally, let $\delta > 0$. Then for $x \in [-1,-\delta]\cup[\delta,1]$, we have that $Q_n(x) = |Q_n(x)| < \sqrt{n}(1-\delta^2)^n$. This shows that $Q_n \to 0$ uniformly as $n \to \infty$ (see proposition 44.D in math 140A notes). So by proposition 120, claim 1.B holds.

Let (Q_n) be the sequence of polynomials in the claim above. Then define $P_n(x) = \int_{-1}^{1} f(x+t)Q_n(t)dt$ for all n.

(Note that P_n is $\mathbb{R}$-valued if f is $\mathbb{R}$-valued)

Claim 2: (P_n) converges uniformly to f on $[0,1]$.

Proof:

Let $\varepsilon > 0$. Then by the uniform continuity of f, there is $\delta \in (0,1)$ such that $\forall s,t \in \mathbb{R},\ \ |s-t| < \delta \implies |f(s)-f(t)| < \frac{\varepsilon}{2}$.

Next set $M = \sup\{|f(x)| \mid x \in (0,1)\}$. By claim 1.B, we can find N so that for all $n \geq N$, $\int_{-1}^{-\delta} Q_n(t)dt + \int_{\delta}^{1} Q_n(t)dt < \frac{\varepsilon}{4M}$.

Now, consider any $x \in [0,1]$ and any $n \geq N$. Then:

$$\begin{aligned}|P_n(x)-f(x)| &= \left|\int_{-1}^{1} f(x+t)Q_n(t)dt - f(x)\cdot(1)\right|\\ &= \left|\int_{-1}^{1} f(x+t)Q_n(t)dt - f(x)\int_{-1}^{1} Q_n(t)dt\right|\\ &= \left|\int_{-1}^{1}(f(x+t)-f(x))Q_n(t)dt\right|\\ &\leq \int_{-1}^{1}|f(x+t)-f(x)|\,Q_n(t)dt\\ &\leq 2M\int_{-1}^{-\delta} Q_n(t)dt + \frac{\varepsilon}{2}\int_{-\delta}^{\delta} Q_n(t)dt + 2M\int_{\delta}^{1} Q_n(t)dt\\ &\leq \frac{\varepsilon}{2}\int_{-1}^{1} Q_n(t)dt + 2M\left(\int_{-1}^{-\delta} Q_n(t)dt + \int_{\delta}^{1} Q_n(t)dt\right)\\ &< \frac{\varepsilon}{2}\cdot 1 + \frac{2M}{4M}\varepsilon = \varepsilon\end{aligned}$$

Claim 3: Each P_n is a polynomial.

Proof:

Note that $P_n(x) = \int_{-1}^{1} f(x+t)Q_n(t)dt = \int_{-x}^{1-x} f(x+t)Q_n(t)dt$. Then, substituting in $u = x+t$, this becomes $P_n(x) = \int_0^1 f(u)Q_n(u-x)du$.

Since $Q_n(u-x)$ is a polynomial, there are finitely many constants $a_{k,l}$ such that $Q_n(u-x) = \sum_{k,l} a_{k,l} u^k x^l$. Therefore:

$$\int_0^1 f(u)Q_n(u-x)du = \int_0^1 \sum_{k,l} a_{k,l} f(u) u^k x^l du = \sum_{k,l} \left(\int_0^1 f(u) u^k du \right) a_{k,l} x^l$$

And now the theorem is proved. ∎

Corollary: For every $a > 0$, there is a sequence (P_n) of real polynomials that converge uniformly to $|x|$ on $[-a, a]$ such that $\forall n, \quad P_n(0) = 0$.

Proof:
By the previous proposition, there is a sequence (P_n^*) of real polynomials that converge to $|x|$ uniformly on $[-a, a]$. In particular, $P_n^*(0) \to 0$. So, set $P_n(x) = P_n^*(x) - P_n^*(0)$.

A collection of functions $\mathscr{A}$ from E to $\mathbb{C}$ is an algebra if:

1. $f, g \in \mathscr{A} \implies f + g \in \mathscr{A}$
2. $f, g \in \mathscr{A} \implies fg \in \mathscr{A}$
3. $f \in \mathscr{A}$ and $c \in \mathbb{C} \implies cf \in \mathscr{A}$

In the case we are discussing $\mathbb{R}$-valued functions, we can change requirement 3 to only consider $c \in \mathbb{R}$.

If $f \in \mathscr{A}$ whenever there is a sequence $(f_n) \subset \mathscr{A}$ converging to f uniformly on E, we say $\mathscr{A}$ is uniformly closed. Also, the set $\mathscr{B}$ consisting of all functions that are uniform limits of functions on E from $\mathscr{A}$ is called the uniform closure of $\mathscr{A}$.

Lecture 18: 5/6/2024

A List of How The Proposition Numbering in my Notes Lines up With Our Textbook:

Proposition Number	Label in Textbook	Proposition Number	Label in Textbook
83	5.2	84	5.3
85	5.5	86	5.8
87	5.9	88	5.10
89	5.11	90	5.12
91	5.13	92	5.15
93	5.19	94	6.4
95	6.5	96	6.6
97	6.7	98	6.8
99	6.9	100	6.10
101	6.11	102	6.12
103	6.13	104	6.15
105	6.16	106	6.17
107	6.19	108	6.20
109	6.21	110	6.22
111	6.25	112	6.27
113	7.8	114	7.9
115	7.10	116	7.11
117	7.12	118	7.13
119	7.15	120	7.16
121	7.17	122	7.18
123	7.23	124	7.24
125	7.25	126	7.26
127		128	
129		130	
131		132	
133		134	

Our textbook is *Principles of Mathematical Analysis* by Walter Rudin.

Homework 1:

Exercise 5.2: Let $f : (a, b) \longrightarrow \mathbb{R}$ be differentiable with $f'(x) > 0$. Then f is strictly increasing.

For all $a < x_1 < x_2 < b$, we know by the mean value theorem that there exists $t \in (x_1, x_2)$ with $f(x_2) - f(x_1) = (x_2 - x_1)f'(t)$. Since $(x_2 - x_1)$ and $f'(t)$ are positive, we thus have that $f(x_2) - f(x_1) > 0$.

As a consequence of f being strictly increasing, we know f is injective. Thus, if we restrict the codomain of f to its image, then f is bijective, meaning there exists a function $g = f^{-1}$ such that $(g \circ f)(x) = x = (f \circ g)(x)$. Now we show that g is differentiable at $f(x)$ for all $x \in \mathrm{dom}(f)$.

Fix $x \in \mathrm{dom}(f)$. Then letting $\varepsilon > 0$, $x_1 = \max(a, x - \varepsilon)$, and $x_2 = \min(b, x + \varepsilon)$, define $c = \inf\limits_{x_1 < t < x_2} f(t)$ and $d = \sup\limits_{x_1 < t < x_2} f(t)$.

Now suppose $s \in (a, b)$ such that $s \leq x_1$. Then because f is strictly increasing, we have that $f(s) < f(t)$ for all $t \in (x_1, x_2)$. Hence, $f(s) \leq c$. Similarly, if $s \geq x_2$, then f being strictly increasing means that $f(s) > f(t)$ for all $t \in (x_1, x_2)$. That in turn would mean that $f(s) \geq d$. So, we've proven by contrapositive that:

$$f(s) \in (c, d) \Longrightarrow s \in (x_1, x_2)$$

Meanwhile because x can't equal a or b we know that $x_1 < x < x_2$. So pick t_1 and t_2 such that $x_1 < t_1 < x < t_2 < x_2$. Then by the definition of supremums and infimums and because f is strictly increasing, we know that $c \leq f(t_1) < f(t_2) \leq d$. Also, because $[t_1, t_2]$ is a connected subset of $\mathrm{dom}(f)$ and f is continuous, we know that at least the connected interval $[f(t_1), f(t_2)] \subseteq [c, d]$ is a subset of $\mathrm{dom}(g)$. At the same time, also because f is strictly increasing, $f(t_1) < f(x) < f(t_2)$.

Therefore, set $\delta = \min(f(x) - f(t_1),\ \ f(t_2) - f(x))$. Then firstly, because $B_\delta(f(x)) \subset \mathrm{dom}(g)$, and $f(x) \in B_\delta(f(x))'$, we know that $f(x)$ is a limit point of $\mathrm{dom}(g)$. Secondly, for any $z \in \mathrm{dom}(g)$, we have that:

$$z = f(s) \in B_\delta(f(x)) \subseteq (c, d) \Longrightarrow g(z) = s \in (x_1, x_2) \subseteq B_\varepsilon(x).$$

Hence, $g(z) \to x$ as $z \to f(x)$.

Finally, consider the limit: $\lim\limits_{z \to f(x)} \frac{g(z) - g(f(x))}{z - f(x)}$ which we can rewrite as $\lim\limits_{z \to f(x)} \frac{g(z) - x}{f(g(z)) - f(x)}$.

Since $f'(x) \neq 0$ for all $x \in \mathrm{dom}(f)$, we can evaluate that $\lim\limits_{t \to x} \frac{t - x}{f(t) - f(x)} = \frac{1}{f'(x)}$. So, given any sequence $(t_n) \subset \mathrm{dom}(f)$ such that $t_n \to x$ and $t_n \neq x$ for any n, we have that:

$$\frac{t_n - x}{f(t_n) - f(x)} \to \frac{1}{f'(g(y))} \text{ as } n \to \infty.$$

Meanwhile, given any sequence $(z_n) \subset \mathrm{dom}(g)$ such that $z_n \to f(x)$ and $z_n \neq f(x)$ for all n, because g is injective and $g(z) \to x$ as $z \to f(x)$, we know that $(g(z_n)) \to x$ as $n \to \infty$ and $g(z_n) \neq x$ for all n.

So for all relevant sequences (z_n), we have that $\frac{g(z_n)-x}{f(g(z_n))-f(x)} \to \frac{1}{f'(x)}$. Hence, $g'(f(x))$ exists with:

$$g'(f(x)) = \lim_{z \to f(x)} \frac{g(z)-g(f(x))}{z-f(x)} = \lim_{z \to f(x)} \frac{g(z)-x}{f(g(z))-f(x)} = \frac{1}{f'(x)}$$

Exercise 5.4: If $C_0 + \frac{C_1}{2} + \ldots + \frac{C_{n-1}}{n} + \frac{C_n}{n+1} = 0$ and $C_0, C_1, \ldots, C_n \in \mathbb{R}$, then we shall prove that the equation $C_0 + C_1 x + \ldots + C_{n-1}x^{n-1} + C_n x^n = 0$ has at least one real root between 0 and 1.

Define the functions:

$$f(x) = C_0 + C_1 x + \ldots + C_{n-1}x^{n-1} + C_n x^n$$
$$F(x) = C_0 x + \frac{C_1}{2}x + \ldots + \frac{C_{n-1}}{n}x^n + \frac{C_n}{n+1}x^{n+1}$$

Note that $F(0) = 0$ and $F(1) = C_0 + \frac{C_1}{2} + \ldots + \frac{C_{n-1}}{n} + \frac{C_n}{n+1} = 0$ At the same time, F is differentiable with $F'(x) = f(x)$. Therefore, by the mean value theorem there exists $t \in (0,1)$ such that $0 = F'(t) = f(t)$. Thus, that t is a real root between 0 and 1 for the equation $C_0 + C_1 x + \ldots + C_{n-1}x^{n-1} + C_n x^n = 0$.

Exercise 5.6: Suppose the following conditions on f:

(A) f is continuous for $x \geq 0$

(B) f' exists for $x > 0$

(C) $f(0) = 0$

(D) f' is monotonically increasing

Putting $g(x) = \frac{f(x)}{x}$ for $x > 0$, we shall prove that g is monotonically increasing.

Firstly, given any $x > 0$, because of conditions A and B, we can apply the mean value theorem to say that there exists $t \in (0, x)$ such that $f(x) - f(0) = xf'(t)$. Because of condition C, this then simplifies to $f(x) = xf'(t)$. So:

for all $x > 0$, there exists $0 < t < x$ such that $\frac{f(x)}{x} = f'(t)$.

Meanwhile, because of condition B and the quotient rule, g is differentiable when $x > 0$ with $g'(x) = \frac{f'(x)x - f(x)}{x^2}$. So, consider any $b > a > 0$. By the mean value theorem, there exists $s \in (a, b)$ with $g(b) - g(a) = (b-a)g'(s)$. Obviously, $b - a$ is positive. Additionally, consider that:

$$g'(s) = \frac{f'(s)s - f(s)}{s^2} = \frac{1}{s}\left(f'(s) - \frac{f(s)}{s}\right).$$

Pick $t > 0$ such that $t < s$ and $\frac{f(s)}{s} = f'(t)$. Then $g'(s) = \frac{1}{s}(f'(s) - f'(t))$. But, because of condition D, we know that $f'(s) \geq f'(t)$. Hence, $g'(s) \geq 0$.

Therefore, $g(b) - g(a) \geq 0$, meaning g is monotonically increasing.

Exercise 5.8: Consider any real-valued function f which is differentiable on $[a,b]$ with f' being continuous on $[a,b]$. Then we shall prove that:

$$\forall \varepsilon > 0, \ \ \exists \delta > 0 \ \ s.t. \ \ \forall x, t \in [a,b], \ \ 0 < |t-x| < \delta \Longrightarrow \left| \frac{f(t)-f(x)}{t-x} - f'(x) \right| < \varepsilon$$

Because f' is continuous over a compact domain, we know that by theorem 4.19 (proposition 76), f' is uniformly continuous. Thus, let $\varepsilon > 0$ and pick $\delta > 0$ such that for all $x, y \in [a,b]$, we have that $|x-y| < \delta \Longrightarrow |f'(x) - f'(y)| < \varepsilon$.

Since f is differentiable on $[a,b]$, we know by the mean value theorem that for any distinct x and t in $[a,b]$, there exists s between a and b such that:

$$\frac{f(t)-f(x)}{t-x} = f'(s).$$

Hence, $\left| \frac{f(t)-f(x)}{t-x} - f'(x) \right| = |f'(s) - f'(x)|$. And since $|s-x| < |t-x|$, we know that if $0 < |t-x| < \delta$, then $|f'(s) - f'(x)| < \varepsilon$.

An analogous theorem holds for any vector-valued function $\vec{f} : [a,b] \longrightarrow \mathbb{R}^k$ that is differentiable on $[a,b]$ with $\vec{f}\,'$ being continuous on $[a,b]$.

Let $\vec{f}(x) = (f_1(x), f_2(x), \ldots, f_k(x))$. Since $\vec{f}$ is differentiable on $[a,b]$ and $\vec{f}\,'$ is continuous on $[a,b]$, we have for each $i \in \{1, \ldots, k\}$ that f_i is differentiable on $[a,b]$ and f_i' is continuous on $[a,b]$.

Thus, given any $\varepsilon > 0$, we already proved that for each $i \in \{1, \ldots, k\}$, there exists $\delta_i > 0$ such that $\forall t, x \in [a,b], \ \ |t-x| < \delta_i \Longrightarrow \left| \frac{f_i(t)-f_i(x)}{t-x} - f_i'(x) \right| < \frac{1}{\sqrt{k}} \cdot \varepsilon$.
Then setting $\delta = \min(\delta_1, \ldots, \delta_k)$, we have that if $0 < |t-x| < \delta$, then:

$$\begin{aligned} \left\| \frac{\vec{f}(t)-\vec{f}(x)}{t-x} - \vec{f}\,'(x) \right\| &= \left(\left(\left| \frac{f_1(t)-f_1(x)}{t-x} - f_1'(x) \right| \right)^2 + \ldots + \left(\left| \frac{f_k(t)-f_k(x)}{t-x} - f_k'(x) \right| \right)^2 \right)^{\frac{1}{2}} \\ &< \left(\left(\frac{1}{\sqrt{k}} \cdot \varepsilon \right)^2 + \ldots + \left(\frac{1}{\sqrt{k}} \cdot \varepsilon \right)^2 \right)^{\frac{1}{2}} = \sqrt{k \left(\frac{1}{k} \cdot \varepsilon^2 \right)} = \varepsilon \end{aligned}$$

Exercise 5.9: Let $x_0 \in (a,b)$ and $f : (a,b) \longrightarrow \mathbb{R}$ be continuous at x_0. If $f'(x)$ exists for all $x \in (a,b) \setminus \{x_0\}$ and $\lim\limits_{t \to x_0} f'(t) = L$, then $f'(x_0) = L$.

Since f is continuous at x_0 and x_0 is a limit point of (a,b), we know that $f(x_0)$ exists and that $\lim\limits_{t \to x_0} f(t) = f(x_0)$. So, define $g(x) = f(x) - f(x_0)$. Then $g'(x) = f'(x)$ and $\lim\limits_{t \to x_0} g(t) = 0$. Additionally, define $h(x) = x - x_0$. Then $h'(x) = 1$ and $\lim\limits_{t \to x_0} h(t) = 0$.

Importantly, both g and h are differentiable everywhere on $(a,b) \setminus \{x_0\}$. Also, $h'(t) \neq 0$ for all $t \in (a,b)$. Thus, we can apply L'hôpital's rule to get that:

$$\lim_{t \to x_0} \frac{f(t)-f(x_0)}{t-x_0} = \lim_{t \to x_0} \frac{g(t)}{h(t)} = \lim_{t \to x_0} \frac{g'(t)}{h'(t)} = \lim_{t \to x_0} f'(t) = L$$

Hence $f'(x_0)$ exists and equals L.

To answer what's actually asked in the book, set $a = -\infty$, $b = +\infty$, $x_0 = 0$ and $L = 3$.

Exercise 5.17: Suppose f is a real, three times differentiable function on $[-1, 1]$ such that $f(-1) = 0$, $f(0) = 0$, $f(1) = 1$, and $f'(0) = 0$. Then $f'''(x) \geq 3$ for some $x \in (-1, 1)$.

Since f is three times differentiable on $[-1, 1]$, we know that f'' is continuous on $[-1, 1]$ and that $f'''(t)$ exists for every $t \in (-1, 1)$. So define:

$$P(t) = f(0) + f'(0)t + \frac{f''(0)}{2}t^2 = \frac{f''(0)}{2}t^2$$

Then by Taylor's theorem, we know that there exists $s \in (0, 1)$ such that $f(1) = P(1) + \frac{f'''(s)}{6}x^3 = \frac{f''(0)}{2} + \frac{f'''(s)}{6}$. Similarly, we know that there exists $t \in (-1, 0)$ such that $f(-1) = \frac{f''(0)}{2} - \frac{f'''(t)}{6}$.

Thus, $\frac{f'''(s)}{6} + \frac{f'''(t)}{6} = f(1) - f(-1) = 1$, which in turn means that $f'''(s) + f'''(t) = 6$. If both $f'''(s)$ and $f'''(t)$ are less than 3, then this is impossible. So, either s or t must be greater than or equal to 3.

Exercise 5.26: Suppose f is differentiable on $[a, b]$, $f(a) = 0$, and there is a real number A such that $|f'(x)| \leq A|f(x)|$ for $x \in [a, b]$. Then $f(x) = 0$ for all $x \in [a, b]$.

To start off, note that if $A < 0$, then we automatically have that $f'(x) = f(x) = 0$ for all $x \in [a, b]$. Meanwhile, if $A = 0$, then $f'(x) = 0$ for all $x \in [a, b]$, thus forcing f to be a constant function. Then, as $f(a) = 0$, we have that $f(x) = f(a) = 0$ for all $x \in [a, b]$.

Therefore, we now assume $A > 0$ and observe the following:

Assume $\gamma \in [a, b)$ and $f(\gamma) = 0$. Then let $x_0 \in [\gamma, b]$ and set $M = \sup_{\gamma \leq x \leq x_0} |f(x)|$.

Then for any $x \in (\gamma, x_0]$, we know by the mean value theorem that there exists $t \in (\gamma, x)$ such that $f(x) - f(\gamma) = (x - \gamma)f'(t)$. Since $f(\gamma) = 0$ and $x > \gamma$, we thus know that $|f(x)| = (x - \gamma)|f'(t)|$. Hence:

$$|f(x)| = (x - \gamma)|f'(t)| \leq (x - \gamma)A|f(t)| \leq A(x - \gamma)M \leq A(x_0 - \gamma)M$$

Now importantly, since f is continuous on $[\gamma, x_0]$, and $g(x) = |x|$ is continuous on all of $\mathbb{R}$, we know that $(g \circ f)(x) = |f(x)|$ is continuous on $[\gamma, x_0]$. That combined with the fact that $[\gamma, x_0]$ is compact means that we can fix $x \in [\gamma, x_0]$ such that $|f(x)| = M$. Then:

- If $x = \gamma$, then $M = |f(\gamma)| = 0$.

- If $x \neq \gamma$, then $M = |f(x)| \leq A(x_0 - \gamma)M$. Crucially, if $\gamma < x_0 < \gamma + \frac{1}{A}$ then $0 < A(x_0 - \gamma) < 1$. Therefore, the only way for $M \leq A(x_0 - \gamma)M$ is if $M = 0$.

Thus, for $x_0 \in [\gamma, \gamma + \frac{1}{A}) \cap [\gamma, b]$, we have that $\sup\limits_{\gamma \leq x \leq x_0} |f(x)| = 0$.

Or in other words, $f(x) = 0$ for all $x \in [\gamma, \gamma + \frac{1}{A}) \cap [\gamma, b]$.

Still assuming $A > 0$, we have that $0 < \frac{1}{2A} < \frac{1}{A}$. So for any $\gamma \in [a, b]$, we know that $[\gamma, \gamma + \frac{1}{2A}) \cap [\gamma, b] \subseteq [\gamma, \gamma + \frac{1}{A}) \cap [\gamma, b]$. Hence, we now proceed by the following inductive process:

Start with $\gamma_1 = a$.
Now do this until told to stop.
If $\gamma_i = b$, then stop. Otherwise, use the above reasoning to show that $f(x) = 0$ for all $x \in [\gamma_i, \min(\gamma_i + \frac{1}{2A}, b)]$. Then set $\gamma_{i+1} = \min(\gamma_i + \frac{1}{2A}, b)$ and repeat these steps with γ_{i+1}.

This process will terminate after $\left\lceil \frac{b-a}{\frac{1}{A}} \right\rceil$ iterations, thus showing that $f(x) = 0$ for all $x \in [a, b]$.

Homework 2:

Exercise 5.11: Suppose f' exists in a neighborhood of x and $f''(x)$ exists. Then:

$$\lim_{h\to 0}\frac{f(x+h)+f(x-h)-2f(x)}{h^2} = f''(x)$$

Define $F(h) = f(x+h)+f(x-h)-2f(x)$ and $G(h) = h^2$. Clearly, both $F(0)=0$ and $G(0)=0$. Plus, because f' exists on a open interval around x, we know that $F'(h)$ exists on an open interval around 0 with $F'(h) = f'(x+h) - f'(x-h)$. At the same time, G' is defined everywhere with $G'(h) = 2h$. Plus, $G'(h) \neq 0$ for any h except $h=0$. So, putting this all together, we can apply L'hopital's rule to get that:

$$\lim_{h\to 0}\frac{f(x+h)+f(x-h)-2f(x)}{h^2} = \lim_{h\to 0}\frac{f'(x+h)-f'(x-h)}{2h} \qquad \text{(assuming the right limit exists)}$$

Meanwhile, note that: $\frac{f'(x+h)-f'(x-h)}{2h} = \frac{1}{2}\left(\frac{f'(x+h)-f'(x)}{h} + \frac{f'(x)-f'(x-h)}{h}\right)$.

Obviously, $\lim\limits_{h\to 0}\frac{f'(x+h)-f'(x)}{h} = f''(x)$. Also, setting $k=-h$, we can say that:

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

So, $\lim\limits_{h\to 0}\frac{f'(x+h)-f'(x-h)}{2h} = \frac{1}{2}(f''(x)+f''(x)) = f''(x)$.

Interestingly, $\lim\limits_{h\to 0}\frac{f(x+h)+f(x-h)-2f(x)}{h^2}$ can be defined even when $f''(x)$ isn't.

A simple example of this is when $f(x) = \begin{cases} 1 & \text{if } x>0 \\ 0 & \text{if } x=0 \\ -1 & \text{if } x<0 \end{cases}$ and we are calculating the limit at $x=0$.

Repeatedly using L'Hôpital's rule, we get that:

$$\lim_{h\to 0}\frac{f(h)+f(-h)-2f(0)}{h^2} = \lim_{h\to 0}\frac{f'(h)-f'(-h)}{2h} = \lim_{h\to 0}\frac{f''(h)+f''(-h)}{2} = \frac{0+0}{2} = 0$$

However, as $f'(0)$ is not defined, obviously $f''(0)$ is not defined either.

For a more interesting example, consider the function $f(x) = \begin{cases} x^2\sin(\frac{1}{x}) & \text{if } x\neq 0 \\ 0 & \text{if } x=0 \end{cases}$ around $x=0$.

As discussed in lecture, $f'(0)$ exists and equals 0. However, f' is discontinuous at 0, which means that $f''(0)$ doesn't exist.

Meanwhile, consider that $f(-x) = -f(x)$. Thus, $f(0+h)+f(0-h)$ cancel out and we have that $f(x+h)+f(x-h)-2f(x) \to 0$ as $h\to 0$. Thus, applying L'Hôpital's rule, we get that:

$$\lim_{h\to 0}\frac{f(h)+f(-h)-2f(0)}{h^2} = \lim_{h\to 0}\frac{f'(h)-f'(-h)}{2h}$$

Next, remember from lecture that $f'(x) = 2x\sin(\frac{1}{x}) - \cos(\frac{1}{x})$ for all $x \neq 0$. Therefore, $f'(-x) = f'(x)$ for all x, which in turn means that $f'(h) - f'(-h) = 0$ for all h again. So, now using L'Hopital's rule again, we get that:

$$\lim_{h\to 0} \frac{f'(h)-f'(-h)}{2h} = \lim_{h\to 0} \frac{f''(h)+f''(-h)}{2}$$

You can check with chain rule that $f''(x) = 2\left(\sin(\frac{1}{x}) - \frac{\cos(\frac{1}{x})}{x}\right) - \frac{\sin(\frac{1}{x})}{x^2}$ for all $x \neq 0$. Also, you can check that $f''(-x) = -f''(x)$ for all x. Thus, $f''(h) + f''(-h) = 0$ for all h, which in turn means that:

$$\lim_{h\to 0} \frac{f''(h)-f''(-h)}{2} = \frac{0}{2} = 0$$

I realize we haven't officially covered sine and cosine yet but in all fairness the professor did bring up this function in class first.

Exercise 5.15: Suppose that $a \in \mathbb{R}$, that f is a twice differentiable real function on (a,∞), and that M_0, M_1, M_2 are the least upper bounds of $|f(x)|, |f'(x)|, |f''(x)|$ on (a,∞) respectively. Then $M_1^2 \leq 4M_0M_2$.

To start, note that if $M_0 = \infty$ or $M_2 = \infty$, then the inequality is trivially true. Also, if M_1 or M_2 equals 0, then we must have $f'(x) = 0$ for all $x \in (a,\infty)$. Hence, the inequality is also true in that case. So, we only need to address when $0 < M_0 < \infty,\ \ 0 < M_1$, and $M_2 < \infty$.

Now, consider any $x \in (a,\infty)$ and $h > 0$. By Taylor's theorem, there exists $\xi \in (x, x+2h)$ such that:

$$f(x+2h) = f(x) + 2h \cdot f'(x) + \frac{f''(\xi)}{2} \cdot 4h^2$$

Or in other words:

$$\begin{aligned} f'(x) &= \tfrac{1}{2h}(f(x+2h) - f(x)) - f''(\xi) \cdot h \\ &\leq \tfrac{1}{2h}(M_0 + M_0) + hM_2 = \tfrac{1}{h}M_0 + hM_2 \end{aligned}$$

If $M_2 = 0$, then we must have that f' is constant, which means that $f'(x) = M_1$ and $M_1 \leq \frac{1}{h}M_0$ for all $h > 0$. But now we must have that $M_0 = \infty$ because otherwise we could pick a large enough h such that $M_1 > \frac{1}{h}M_0$. So, $M_1^2 \leq 4M_0M_2$ is true again.

If $M_1 = \infty$, then by setting $h = 1$, we must have that $M_0 + M_2$ is greater than all positive real numbers. Thus, because either M_1 or M_2 equals ∞ and M_1 and M_2 can't equal $-\infty$, we have that $M_1^2 \leq 4M_0M_2$ is true trivially.

So, now that we can finally assume M_0, M_1, and M_2 are all finite and nonzero, set $h = \sqrt{\frac{M_0}{M_2}}$. Then $\frac{1}{h}M_0 + hM_2 = 2\sqrt{M_0}\sqrt{M_2}$. So, $(f'(x))^2 = |f'(x)|^2 \leq 4M_0M_2$ for all $x \in (a,\infty)$. Or in other words, $M_1^2 \leq 4M_0M_2$.

It's also possible for $M_1^2 = 4M_0M_2$.

Set $f(x) = \begin{cases} 2x^2 - 1 & \text{if } -1 < x < 0 \\ \frac{x^2-1}{x^2+1} & \text{if } x \geq 0 \end{cases}$

Then, $f'(x) = 4x$ when $-1 < x < 0$, whereas $f'(x) = \frac{4x}{(x^2+1)^2}$ when $x > 0$. Importantly, this shows that $\lim_{t \to x} f'(t) = 0$. So, by applying homework exercise 5.9 from last week, we know that $f'(0) = 0$ as well.

Next, $f''(x) = 4$ when $-1 < x < 0$, whereas $f''(x) = \frac{-12x^2+4}{(x^2+1)^3}$ when $x > 0$. Importantly, we can see that $\lim_{t \to x} f''(t) = 4$. So, by applying homework exercise 5.9 from last week again, we know that $f''(0) = 4$ as well.
So:

$$f(x) = \begin{cases} 2x^2 - 1 & \text{if } -1 < x < 0 \\ \frac{x^2-1}{x^2+1} & \text{if } x \geq 0 \end{cases} \qquad f'(x) = \begin{cases} 4x & \text{if } -1 < x < 0 \\ \frac{4x}{(x^2+1)^2} & \text{if } x \geq 0 \end{cases}$$

$$f''(x) = \begin{cases} 4 & \text{if } -1 < x < 0 \\ \frac{-12x^2+4}{(x^2+1)^3} & \text{if } x \geq 0 \end{cases}$$

Then $M_0 = 1$, $M_1 = 4$, and $M_2 = 4$. Check those on your own time because I'm running out of time to turn this in. :P

So $M_1^2 = 4M_0M_2$.

If $\vec{f}$ is a twice differentiable $\mathbb{R}^k$-valued function on (a, ∞) and M_0, M_1, M_2 are the least upper bounds of $\|\vec{f}(x)\|, \|\vec{f}'(x)\|, \|\vec{f}''(x)\|$ on (a, ∞) respectively, then we still have that $M_1^2 \leq 4M_0M_2$.

If $M_1 = 0$, the inequality is true trivially. So let's assume that $M_0 > 0$. Then for any $0 < \alpha < M_1$, we can pick x_0 such that $\|\vec{f}'(x)\| > \alpha$. Next, define the real valued function $g(x) = \frac{1}{\|\vec{f}(x_0)\|} \vec{f}'(x_0) \cdot \vec{f}(x)$.

Note that g is a twice differentiable real function defined on (a, ∞). So, let N_0, N_1, N_2 be the least upper bounds of $|g(x)|, |g'(x)|, |g''(x)|$ on (a, ∞) respectively. By part 1 of this exercise, we know that $N_1^2 \leq 4N_0N_2$.

Also, note that $g'(x) = \frac{1}{\|\vec{f}(x_0)\|} \vec{f}'(x_0) \cdot \vec{f}'(x)$ and $g''(x) = \frac{1}{\|\vec{f}(x_0)\|} \vec{f}'(x_0) \cdot \vec{f}''(x)$.
Thus, by the Cauchy-Schwarz inequality:

$g(x) \leq \|\vec{f}(x)\|$, $g'(x) \leq \|\vec{f}'(x)\|$, and $g''(x) \leq \|\vec{f}''(x)\|$ for all $x \in (a, \infty)$.

Importantly, this means that $N_0 \leq M_0$ and $N_2 \leq M_2$. Therefore, $N_1^2 \leq 4M_0M_2$.

Also, note that $g'(x_0) = \|\vec{f}'(x_0)\| > \alpha$. So, because $\alpha < N_1$, we have that $\alpha \leq 4M_0M_2$. And since α is any positive number less than M_1, we thus have that $M_1 \leq 4M_0M_2$.

Exercise 5.22: Suppose f is a real function on $(-\infty, \infty)$. We call x a fixed point of f if $f(x) = x$. Firstly, we show that if f is differentiable and $f'(t) \neq 1$ for any t, than f has at most one fixed point.

Assume $f(x) = x$ and $f(y) = y$ for some $x, y \in \mathbb{R}$. If $x \neq y$, then by the mean value theorem, there exists $t \in (x, y)$ such that:
$$y - x = f(y) - f(x) = (y - x)f'(t)$$

But since $f'(t) \neq 1$ for any t, this is impossible. So, we conclude that x must equal y.

Secondly, we shall show that $f(t) = t + (1 + e^t)^{-1}$ has no fixed points but that $0 < f'(t) < 1$ for all real t.

Since $\frac{1}{1+e^t} > 0$ for all t, we automatically have that $t < f(t)$ for all t. Hence, f can have no fixed point.

Meanwhile, $f'(t) = 1 - \frac{e^t}{(1+e^t)^2}$. Because $e^t > 0$ and $(1 + e^t)^2 > e^t > 0$, we know that $0 < \frac{e^t}{(1+e^t)^2} < 1$. Hence, $0 < f'(t) < 1$. for all t.

Thirdly, we show that if there is a constant $A < 1$ such that $|f'(t)| \leq A$ for all real t, then f has a fixed point x. Furthermore, $x = \lim x_n$ where $x_1 \in \mathbb{R}$ and $x_{n+1} = f(x_n)$ for all $n \in \mathbb{Z}_+$.

To start off, note that if $x_n = x_{n+1}$ for any value of $n \in \mathbb{Z}_+$, then $x_n = x_{n+k}$ for all $k \in \mathbb{Z}_+$ and $x_n = f(x_n)$. So, we trivially have that $x_n \to x$ where x is the fixed point of f.

Now we assume that $x_n \neq x_{n+1}$ for any $n \in \mathbb{Z}_+$. Then for any $n \in \mathbb{Z}_+$, we can use the mean value theorem to say that there exists $t \in (x_n, x_{n+1})$ such that:
$$x_{n+2} - x_{n+1} = f(x_{n+1}) - f(x_n) = (x_{n+1} - x_n)f'(t) < (x_{n+1} - x_n)A$$

So for any $n \in \mathbb{Z}_+ \setminus \{1\}$, we can say that:
$$|x_{n+1} - x_n| < |x_n - x_{n-1}|A < \ldots < |x_2 - x_1|A^{n-1}.$$

In turn, this means that for all integers $m > n \geq 2$, we have that:
$$|x_m - x_n| \leq \sum_{i=1}^{m-n} |x_{n+i} - x_{n+i-1}| < \sum_{i=1}^{m-n} |x_2 - x_1|A^{n+i-2}$$
$$< |x_2 - x_1|A^{n-1} \sum_{i=0}^{m-n-1} A^{i-1} < \frac{|x_2 - x_1|}{1-A} A^{n-1}$$

Now let $\varepsilon > 0$ and pick N big enough so that $A^{N-1} < \frac{\varepsilon(1-A)}{|x_2 - x_1|}$. Then for all $m > n > N$, we have $|x_m - x_n| < \frac{|x_2-x_1|}{1-A} A^{n-1} < \frac{|x_2-x_1|}{1-A} A^{N-1} < \varepsilon$. Hence, we have shown that (x_n) is Cauchy. And since $\mathbb{R}$ is complete, we thus have that (x_n) converges.

Let x be the limit of (x_n) as n goes to ∞. Then let $\varepsilon > 0$. Since f is differentiable at x, we know that f is continuous at x. So, there exists $\delta > 0$ such that:

$$|x - x_m| < \delta \implies |f(x) - f(x_m)| < \varepsilon/2.$$

Meanwhile, since $x_n \to x$, there exists $N \in \mathbb{Z}_+$ such that:

$$n > N \implies |x_n - x| < \min(\delta, \varepsilon/2).$$

So, pick an integer $m > N$. Then:

$$|f(x) - x| \leq |f(x) - f(x_m)| + |x_{m+1} - x| < \varepsilon/2 + \varepsilon/2 < \varepsilon$$

Hence, $f(x) = x$.

One way to visualize this is by a zig-zag path in $\mathbb{R}^2$:

$$(x_1, x_2) \to (x_2, x_2) \to (x_2, x_3) \to \ldots$$

That sequence of ordered pairs converges because each individual coordinate converges. I don't know what else Rudin expects me to say.

Exercise 6.1: Suppose that α increases on $[a, b]$, that $a \leq s \leq b$, that α is continuous at s, that $f(s) = 1$, and that $f(x) = 0$ if $x \neq s$. Then, we shall show that $f \in \mathscr{R}_a^b(\alpha)$ and that $\int_a^b f d\alpha = 0$.

To start off, we know by theorem 6.10 (proposition 100) that $f \in \mathscr{R}_a^b(\alpha)$. After all, f is bounded and has only one discontinuity, and α is continuous where f is discontinuous. So, let $C = \int_a^b f d\alpha$ and $\varepsilon > 0$. Then, pick a partition $P = \{x_0, \ldots, x_n\}$ of $[a, b]$ such that $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$.

Next, for each $i \in \{1, \ldots, n\}$, pick $t_i \in [x_{i-1}, x_i]$ such that $t_i \neq s$. Then by theorem 6.7C (proposition 97.C), we know that:

$$\left| \sum_{i=1}^{n} f(t_i) \Delta\alpha_i - c \right| \leq U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

Now importantly, $f(t_i) = 0$ since $t_i \neq s$. Hence, we've just shown that $|c| < \varepsilon$. And since ε was arbitrary, we know that $\int_a^b f d\alpha = c = 0$.

Exercise 6.2: Let $f : [a, b] \longrightarrow \mathbb{R}$ such that $f(x) \geq 0$ and f is continuous at x for all $x \in [a, b]$. Then given that $\int_a^b f(t)dt = 0$, we show that $f(x) = 0$ for all $x \in [a, b]$.

Suppose $f(s) \neq 0$ for some $s \in [a, b]$. Then, because f is continuous at s, there exists $\delta > 0$ such that $|x - s| < \delta_1 \implies |f(x) - f(s)| < \frac{f(s)}{2}$. Or in other words, $|x - s| < \delta_1 \implies \frac{f(s)}{2} < f(x)$.

Now set $\delta_2 = \min\{\frac{\delta_1}{2}, \frac{b-a}{2}\}$ and pick a partition $P = \{x_0, \ldots, x_n\}$ of $[a, b]$ such that $s \in [x_{j-1}, x_j]$ for some $j \in 1, \ldots, n$, and that $x_j - x_{j-1} = \delta_2$. Then for the jth interval of P, we have that $m_j \geq \frac{f(s)}{2}$. So:

$$L(P, f) = m_j \Delta x_j + \sum_{\substack{i=1 \\ i \neq j}}^{n} m_i \Delta x_i > m_j \Delta x_j = m_j \delta_2 \geq \frac{f(s)}{2} \delta_2 > 0$$

So, we now have a contradiction because $\int_a^b f(t)dt = \underline{\int_a^b} f(t)dt \geq \frac{f(s)}{2}\delta_2 > 0$.

Hence, we conclude that there cannot be a point s where $f(s) \neq 0$.

Exercise 6.4: Let $f(x) = 0$ if $x \in \mathbb{R} \setminus \mathbb{Q}$, and let $f(x) = 1$ if $x \in \mathbb{Q}$. Then $f \notin \mathscr{R}_a^b$ for any $a < b$.

Let P be any partition of $[a, b]$. Then for any subinterval, we have that $m = 0$ and $M = 1$. Hence, $U(P, f) = b - a$ and $L(P, f) = 0$. This means that $\underline{\int_a^b} f(t)dt = 0$ and $\overline{\int_a^b} f(t)dt = b - a$. So as $b \neq a$, we have that $\underline{\int_a^b} f(t)dt \neq \overline{\int_a^b} f(t)dt$, meaning that $\int_a^b f(t)$ is not defined.

Homework 3:

Exercise 6.5: Even if f is a bounded real function on $[a,b]$ and $f^2 \in \mathscr{R}_a^b$, we can't guarentee that $f \in \mathscr{R}_a^b$.

Define $f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \\ -1 & \text{if } x \in \mathbb{Q} \end{cases}$

Clearly, f is bounded. Also, $f^2(x) = 1$, which means that $f^2 \in \mathscr{R}_a^b$. But, for any partition P of $[a,b]$, we have that $L(P,f) = -(b-a)$ and $U(P,f) = +(b-a)$. So, $\underline{\int_a^b} f\,dx \neq \overline{\int_a^b} f\,dx$ when $b \neq a$, which mean that $f \notin \mathscr{R}_a^b$.

However, if f is a bounded real function on $[a,b]$ and $f^3 \in \mathscr{R}_a^b$, then we can guarentee that $f \in \mathscr{R}_a^b$.

Define $\phi(x) = \begin{cases} \sqrt[3]{x} & \text{if } x \geq 0 \\ -\sqrt[3]{-x} & \text{if } x < 0 \end{cases}$

(I'm trying to be a little careful because we only proved the existence of roots of positive numbers in chapter 1)

Now firstly we shall show that ϕ is continuous.

If $x, t > 0$, then note that $\left(\sqrt[3]{t} + \sqrt[3]{x}\right)^2 = \sqrt[3]{t^2} + 2\sqrt[3]{tx} + \sqrt[3]{x^2} \neq 0$. So:

$$\sqrt[3]{t} - \sqrt[3]{x} = \left(\sqrt[3]{t} - \sqrt[3]{x}\right)\frac{\sqrt[3]{t^2} + 2\sqrt[3]{tx} + \sqrt[3]{x^2}}{\sqrt[3]{t^2} + 2\sqrt[3]{tx} + \sqrt[3]{x^2}} = \frac{t-x}{\sqrt[3]{t^2} + 2\sqrt[3]{tx} + \sqrt[3]{x^2}} \leq \frac{t-x}{\sqrt[3]{tx}}$$

Similarly, if $x, t < 0$, then $\left(\sqrt[3]{-x} + \sqrt[3]{-t}\right)^2 \neq 0$. So:

$$\sqrt[3]{-x} - \sqrt[3]{-t} = \left(\sqrt[3]{-x} - \sqrt[3]{-t}\right)\frac{\sqrt[3]{(-x)^2} + 2\sqrt[3]{-t \cdot -x} + \sqrt[3]{(-t)^2}}{\sqrt[3]{(-x)^2} + 2\sqrt[3]{-t \cdot -x} + \sqrt[3]{(-t)^2}}$$

$$= \frac{-x - -t}{\sqrt[3]{t^2} + 2\sqrt[3]{tx} + \sqrt[3]{x^2}} \leq \frac{t-x}{\sqrt[3]{tx}}$$

Thus, when $x \neq 0$ and $t \neq 0$, we know that $|\phi(t) - \phi(x)| \leq \frac{|t-x|}{|\sqrt[3]{tx}|} = \frac{|t-x|}{\sqrt[3]{|t||x|}}$.

Now, letting $\varepsilon > 0$, we shall try to find δ such that $|t-x| < \delta \implies \frac{|t-x|}{\sqrt[3]{|t||x|}} < \varepsilon$. To do this, assume $|t| > \frac{|x|}{8}$. Then $\frac{|t-x|}{\sqrt[3]{|t||x|}} < \frac{2|t-x|}{x^{2/3}}$. So, setting $\delta = \min(\frac{7|x|}{8}, \frac{\varepsilon x^{2/3}}{2})$, we have that $|t-x| < \delta \implies |\phi(t) - \phi(x)| \leq \frac{|t-x|}{\sqrt[3]{|t||x|}} < \frac{2|t-x|}{x^{2/3}} < \varepsilon$. Hence, ϕ is continuous at $x \neq 0$.

Now let's address when $x = 0$. For any $\varepsilon > 0$, pick $\delta = \varepsilon^3$. Then:

$0 < t < \delta \implies 0 < \sqrt[3]{t} < \varepsilon$ and $-\delta < t < 0 \implies 0 > -\sqrt[3]{-t} > -\varepsilon$.

Hence, ϕ is also continuous at $x = 0$.

Secondly, observe that $f = \phi \circ f^3$. So by theorem 6.11 (proposition 101), we know that $f = \phi \circ f^3 \in \mathscr{R}_a^b$.

Exercise 6.6: Let C be the Cantor set and f be a bounded real function on $[0,1]$ which is continuous at every point outside C. Then $f \in \mathscr{R}_0^1$.

Set $M > 0$ such that $M \geq \sup\{|f(x)| \mid 0 \leq x \leq 1\}$. That way we can divide by M. Then consider any $\varepsilon > 0$.

Now, remember that $C = \bigcap_{i=0}^{\infty} C_i$ where $C_0 = [0,1]$, $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$, ...
Note that the "length" of C_n is $\left(\frac{2}{3}\right)^n$. Obviously, that goes to 0 as n goes to infinity. So, pick $N \in \mathbb{Z}_+$ such that $\left(\frac{2}{3}\right)^{N-1} < \frac{\varepsilon}{4M}$. Then suppose $C_N = \bigcup_{j=1}^{m} [a_j, b_j]$ for some $m \in \mathbb{Z}_+$. Based on that, pick out $u_j < v_j$ for each j such that:

- $a_j - \frac{1}{3^{N+1}} < u_j < a_j$ when $a_j \neq 0$
- $b_j < v_j < b_j + \frac{1}{3^{N+1}}$ when $b_j \neq 1$
- $u_j = 0$ when $a_j = 0$ and $v_j = 1$ when $b_j = 1$

With that done, we now have that $\sum_{j=1}^{m} (v_j - u_j) < \frac{\varepsilon}{4M}$, and that $C \subset \bigcup_{j=1}^{m} (u_j, v_j)$.

Now set $K = [0,1] \setminus \bigcup_{j=1}^{m} (u_j, v_j)$ and notice that K is closed and bounded in $\mathbb{R}$. Hence, K is compact. Also, $K \cap C = \{0, 1\}$. Thus, because 0 and 1 are isolated points of K, we know that f is continous over K. So, we can conclude that f is uniformly continuous over K.

Therefore, there exists $\delta > 0$ such that $\forall s,t \in K, \ \ |s - t| < \delta \implies |f(s) - f(t)| < \varepsilon/2$.
So, pick a partition $P = \{x_0, \ldots, x_n\}$ of $[0,1]$ such that:

- $\{u_1, v_1, u_2, v_2, \ldots, u_m, v_m\} \subseteq P \subseteq K$
- $x_{i-1} \notin \{u_1, u_2, \ldots u_m\} \implies \Delta x_i < \delta$ for all $i \in \{1, \ldots, n\}$.

Then after defining M_i and m_i like normal, note that:

$$(M_i - m_i)\Delta x_i \leq \begin{cases} 2M(v_j - u_j) & \text{if } x_{i-1} = u_j \text{ for some } j \\ \varepsilon/2 \Delta x_i & \text{if } x_{i-1} \notin \{u_1, \ldots, u_m\} \end{cases}$$

So:

$$\begin{aligned} U(P,f) - L(P,f) &= \sum_{i=1}^{n} (M_i - m_i)\Delta x_i \\ &\leq 2M \sum_{j=1}^{m} (v_j - u_j) + \varepsilon/2(1 - 0) + 2M \tfrac{\varepsilon}{4M} + \varepsilon/2 = \varepsilon \end{aligned}$$

Hence, we have shown we can find a partition P such that $U(P,f) - L(P,f) < \varepsilon$.
So, $f \in \mathscr{R}_0^1$.

Exercise 6.7: Suppose that f is a real function on $(0,1]$ and that $f \in \mathscr{R}^1_c$. for every $c > 0$. Also, suppose $I = \lim\limits_{c \to 0} \int_c^1 f(x)dx$ exists.

(A) If $f \in \mathscr{R}^1_0$, then $I = \int_0^1 f(x)dx$.

Let $\varepsilon > 0$.

Since $f \in \mathscr{R}^1_0$, when we define $F(x) = \int_0^x$ on the domain $[0,1]$, we know by theorem 6.20 (proposition 108) that F is continuous. So, $F(c) \to F(0) = 0$ as $c \to 0$. At the same time, note that:

$$\int_c^1 f(x)dx = \int_0^1 f(x)dx - \int_0^c f(x)dx = \int_0^1 f(x)dx - F(c)$$

So $I = \lim\limits_{c \to 0} \int_c^1 f(x)dx = \lim\limits_{c \to 0} \left(\int_0^1 f(x)dx - F(c) \right) = \int_0^1 f(x)dx - 0 = \int_0^1 f(x)dx$.

(B) It's possible for $\lim\limits_{c \to 0} \int_c^1 f(x)dx$ to exist but not $\lim\limits_{c \to 0} \int_c^1 |f(x)|dx$.

We will show this by modeling f off of the series $\sum \frac{(-1)^{n-1}}{n}$. After all, we know that that series converges but not absolutely.

Note that $\left(\frac{1}{n} - \frac{1}{n+1}\right)^{-1} = n(n+1)$. Thus $\left(\frac{1}{n} - \frac{1}{n+1}\right)(n+1) = \frac{1}{n}$ for all $n \in \mathbb{Z}_+$. So, let's now define $f(x) = (-1)^{n-1}(n+1)$ where n is the unique integer such that $\frac{1}{n+1} < x \leq \frac{1}{n}$. Then, if $N \in \mathbb{Z}_+$ such that $\frac{1}{N+1} < c \leq \frac{1}{N}$, we have that:

$$\int_c^1 f(x)dx = (\tfrac{1}{N} - c)(-1)^{N-1}(N+1) + \sum_{k=1}^{N-1} \frac{(-1)^{k-1}}{k}$$

Now $0 \leq (\frac{1}{N} - c)(N+1) < (\frac{1}{N} - \frac{1}{N+1})(N+1) = \frac{1}{N}$. Therefore, we have that $\left|\frac{1}{N} - c)(-1)^{N-1}(N+1)\right| < \frac{1}{|N|}$. Then because $\frac{1}{|N|} \to 0$ as $N \to \infty$ which happens as $c \to 0$, we know that $(\frac{1}{N} - c)(-1)^{N-1}(N+1) \to 0$ as $c \to 0$. Hence:

$$\begin{aligned}\lim_{c \to 0} \int_c^1 f(x)dx &= \lim_{c \to 0} \left((\tfrac{1}{N} - c)(-1)^{N-1}(N+1) + \sum_{k=1}^{N-1} \frac{(-1)^{k-1}}{k} \right) \\ &= 0 + \lim_{c \to 0} \sum_{k=1}^{N-1} \frac{(-1)^{k-1}}{k} = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}\end{aligned}$$

Meanwhile, $\int_c^1 |f(x)|dx = (\frac{1}{N} - c)(N+1) + \sum\limits_{k=1}^{N-1} \frac{1}{k}$. As before, we know that $(\frac{1}{N} - c)(N+1) \to 0$ as $c \to 0$. Hence:

$$\lim_{c \to 0} \int_c^1 |f(x)|dx = \lim_{c \to 0} \left((\tfrac{1}{N} - c)(N+1) + \sum_{k=1}^{N-1} \frac{1}{k} \right) = 0 + \lim_{c \to 0} \sum_{k=1}^{N-1} \frac{1}{k} = \sum_{k=1}^{\infty} \frac{1}{k}$$

Now $\sum\limits_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}$ converges whereas $\sum\limits_{k=1}^{\infty} \frac{1}{k}$ does not converge.

Therefore, $\lim\limits_{c \to 0} \int_c^1 f(x)dx$ exists but not $\lim\limits_{c \to 0} \int_c^1 |f(x)|dx$.

Exercise 6.8: Suppose $f \in \mathscr{R}_a^b$ for every $b > a$ where a is fixed. Then define $\int_a^\infty f(x)dx = \lim\limits_{b\to\infty} \int_a^b f(x)dx$ when that limit exists and is finite. Also, when that limit exists we say that $\int_a^\infty f(x)dx$ "converges".

Now assume that $f(x) \geq 0$ and that f decreases monotonically on $[1, \infty)$. Then:

$\int_1^\infty f(x)dx$ converges if and only if $\sum\limits_{n=1}^{\infty} f(n)$ converges.

Let $b \in (1, \infty)$ and set $N \in \mathbb{Z}_+$ such that $N < b \leq N+1$. Then consider the partition $P = \{1, 2, \ldots, N-1, N, b\}$. Importantly:

$L(P, f) = \sum\limits_{k=1}^{N-1} f(k+1) + f(b)(b-N)$ and $U(P, f) = \sum\limits_{k=1}^{N-1} f(k) + f(N)(b-N)$.

Also, note that if $F(x) = \int_1^x f(t)dt$, then F is monotone increasing. Plus, the partial sums of $\sum f(n)$ are also monotone increasing. Hence, $\lim\limits_{b\to\infty} F(b) = \int_1^\infty f(t)dt$ exists if and only if $F(b)$ is bounded and $\sum\limits_{n=1}^{\infty} f(n)$ exists if and only if the partial sums are bounded.

($\Longrightarrow$) Suppose $\int_1^\infty f(x)dx$ converges. Then consider that:

$$L(P, f) = -f(1) + \sum_{k=1}^{N} f(k) + f(b)(b-N) \leq \int_0^b f(x)dx$$

Hence $\sum\limits_{k=1}^{N} f(k) \leq \int_1^b f(x)dx + f(1) - f(b)(b-N) \leq \int_1^\infty f(x)dx + f(1) - 0.$

This means that the partial sums of $\sum f(n)$ are bounded. So, $\sum\limits_{n=1}^{\infty} f(n)$ converges.

($\Longleftarrow$) Now suppose $\sum\limits_{n=1}^{\infty} f(n)$ converges. Then consider that:

$$F(b) = \int_0^b f(x)dx \leq U(P, f) = \sum_{k=1}^{N-1} f(k) + f(N)(b-N) \leq \sum_{k=1}^{N} f(K) \leq \sum_{n=1}^{\infty} f(n)$$

This shows that F is bounded above by $\sum\limits_{n=1}^{\infty} f(n)$. So, $\lim\limits_{b\to\infty} F(b) = \int_1^\infty f(t)dt$ exists.

Exercise 6.10: Let p and q be positive real numbers such that $\frac{1}{p}+\frac{1}{q}=1$.

(A) If $u \geq 0$ and $v \geq 0$, then $uv \leq \frac{u^p}{p}+\frac{v^q}{q}$ with equality holding if and only if $u^p = v^q$.

Firstly, note that because neither p nor q are less than or equal to 0 and $\frac{1}{p}+\frac{1}{q}=1$, we know that both $p>1$ and $q>1$.

Now fix $v \geq 0$ and define $f(u)=\frac{u^p}{p}+\frac{v^q}{q}-uv$. Putting aside that we haven't yet shown that the power rule applies to non positive integer exponents, note that $f'(u)=u^{p-1}-v$ exists for all $u \geq 0$.

Importantly, $f'(0)=-v$. Also, because $p-1>0$, we know that u^{p-1} is strictly increasing and unbounded as $u \to \infty$. As a result, using theorem 5.12 (proposition 90) we know there exists $u_0 \in (0,\infty)$ such that $f'(u_0)=0$. Furthermore, $f'(u)<0$ if $u<u_0$ and $f'(u)>0$ if $u>u_0$. In turn this let's us proof using the mean value theorem that f reaches its minimum value over all $u \geq 0$ precisely when $u=u_0$.

Now, $u_0^{p-1}-v=0 \Longrightarrow u_0 = v^{\frac{1}{p-1}}$. Therefore:

$$u_0^p = v^{\frac{p}{p-1}} = v^{\left(1+\frac{1}{p-1}\right)} = v \cdot v^{\frac{1}{\frac{1}{1-\frac{1}{q}}-1}} = v \cdot v^{\frac{1}{\frac{q}{q-1}-1}} = v \cdot v^{\frac{q-1}{q-q+1}} = v \cdot v^{q-1} = v^q$$

Also, $u_0 = v^{\frac{1}{p-1}} = v^{\frac{1}{\frac{1}{1-\frac{1}{q}}-1}} = v^{\frac{1}{\frac{q}{q-1}-1}} = v^{\frac{q-1}{q-q+1}} = v^{q-1}$. So:

$$f(u_0) = \frac{v^q}{p}+\frac{v^q}{q} - v^{q-1}\cdot v = v^q\left(\frac{1}{p}+\frac{1}{q}-1\right) = v^q(0)=0.$$

Thus, we have shown that $f(u) \geq 0$ with $f(u)=0$ when $u^p=v^q$. Or in other words, $uv \leq \frac{u^p}{p}+\frac{v^q}{q}$ with equality holding if and only if $u^p=v^q$.

(B) If $f,g \in \mathscr{R}_a^b(\alpha)$, $\quad f,g \geq 0\;$ and $\int_a^b f^p d\alpha = 1 = \int_a^b g^q d\alpha$, then $\int_a^b fg d\alpha \leq 1$.

Note that $f(x)g(x) \leq \frac{f^p(x)}{p}+\frac{g^q(x)}{q}$ by part A. Importantly, both sides of that inequality are integratable via theorems 6.12 and 6.13 (propositions 102 and 103). Therefore, integrating both sides of the inequality, we know by theorem 6.12 that:

$$\int_a^b fg d\alpha \leq \int_a^b \left(\frac{f^p}{p}+\frac{g^q}{q}\right) d\alpha = \frac{1}{p}+\frac{1}{q}=1$$

(C) If $f,g : [a,b] \longrightarrow \mathbb{C}$ and $f,g \in \mathscr{R}_a^b(\alpha)$, then:

$$\left|\int_a^b fg d\alpha\right| \leq \left(\int_a^b |f|^p d\alpha\right)^{\frac{1}{p}} \left(\int_a^b |g|^q d\alpha\right)^{\frac{1}{q}}$$

Suppose $\int_a^b |f|^p d\alpha = 0$. Then because $|f| \geq 0$, we must have that $|f(x)| = 0$ for all $x \in [a, b]$. But then $f(x)$ and in turn $f(x)g(x)$ equal 0 over $[a, b]$, which means that $\int_a^b fgd\alpha = 0$. So, the stated inequality is true trivially.

Similarly, if $\int -a^b |g|^q d\alpha = 0$, then the stated inequality is true trivially. So, it is safe to assume that $\left(\int_a^b |f|^p d\alpha\right)^{\frac{1}{p}}$ and $\left(\int_a^b |g|^q d\alpha\right)^{\frac{1}{q}}$ are strictly greater than 0.

Now define $F(x) = \frac{1}{\left(\int_a^b |f|^p d\alpha\right)^{\frac{1}{p}}}|f(x)|$ and $G(x) = \frac{1}{\left(\int_a^b |g|^q d\alpha\right)^{\frac{1}{q}}}|g(x)|$.

Clearly, $\int_a^b F d\alpha = 1 = \int_a^b G d\alpha$. with $F, G \geq 0$ and $F, G \in \mathscr{R}_a^b$. So by part B, we know that:

$$\int_a^b FGd\alpha = \frac{1}{\left(\int_a^b |f|^p d\alpha\right)^{\frac{1}{p}}} \cdot \frac{1}{\left(\int_a^b |g|^q d\alpha\right)^{\frac{1}{q}}} \cdot \int_a^b |f||g| d\alpha \leq 1$$

Then because $\left|\int_a^b fgd\alpha\right| \leq \int_a^b |fg| d\alpha = \int_a^b |f||g| d\alpha$, we can conclude that:

$$\left|\int_a^b fgd\alpha\right| \leq \left(\int_a^b |f|^p d\alpha\right)^{\frac{1}{p}} \left(\int_a^b |g|^q d\alpha\right)^{\frac{1}{q}}$$

Note: this is called Hölder's inequality.

(D) Hölder's inequality also holds for the "improper integrals described in exercises 6.7 and 6.8.

Firstly, suppose that f and g are real functions on $(a, b]$ and that $f \in \mathscr{R}_c^1$. for every $c > a$. Also, suppose $\lim_{c \to a} \int_c^b f(x)dx$ and $\lim_{c \to a} \int_c^b g(x)dx$ exist.

Now we know that $\left|\int_c^b f(x)g(x)dx\right| \leq \left(\int_c^b |f|^p dx\right)^{\frac{1}{p}} \left(\int_c^b |g|^q dx\right)^{\frac{1}{q}}$.

Thus, we have that $\lim_{c \to a} \left|\int_c^b f(x)g(x)dx\right| \leq \lim_{c \to a} \left(\left(\int_c^b |f(x)|^p dx\right)^{\frac{1}{p}} \left(\int_c^b |g(x)|^q dx\right)^{\frac{1}{q}}\right)$ when the limits of both sides exist.

Now if $\lim_{c \to a} \int_c^b f(x)g(x)dx$, $\lim_{c \to a} \int_c^b |f(x)|^p dx$, and $\lim_{c \to a} \int_c^b |g(x)|^p dx$ exist, then we know that $\lim_{c \to a} \left|\int_c^b f(x)g(x)dx\right| = \left|\lim_{c \to a} \int_c^b f(x)g(x)dx\right|$ and $\lim_{c \to a} \left(\left(\int_c^b |f(x)|^p dx\right)^{\frac{1}{p}} \left(\int_c^b |g(x)|^q dx\right)^{\frac{1}{q}}\right) = \left(\lim_{c \to a} \int_c^b |f(x)|^p dx\right)^{\frac{1}{p}} \left(\lim_{c \to a} \int_c^b |g(x)|^q dx\right)^{\frac{1}{q}}$

So $\left|\lim_{c \to a} \int_c^b f(x)g(x)dx\right| \leq \left(\lim_{c \to a} \int_c^b |f(x)|^p dx\right)^{\frac{1}{p}} \left(\lim_{c \to a} \int_c^b |g(x)|^q dx\right)^{\frac{1}{q}}$.

Now while I exclusively dealt with the improper integrals from exercise 6.7 just now, you can do the same reasoning on the improper integrals from exercise 6.8.

Homework 5:

Exercise 7.2: If (f_n) and (g_n) converge uniformly on a set E to f and g respectively, then $(f_n + g_n)$ converges uniformly on E.

Since both (f_n) and (g_n) converge uniformly, for any $\varepsilon > 0$ we can pick N_1 and N_2 such that:

- $\forall x \in E, \ \ \forall n \geq N_1, \ \ |f_n(x) - f(x)| < \varepsilon/2$
- $\forall x \in E, \ \ \forall n \geq N_2, \ \ |g_n(x) - g(x)| < \varepsilon/2$

Now set $N = \max(N_1, N_2)$. Then for all $n \geq N$ and $x \in E$, we have that:

$$|(f_n + g_n)(x) - (f+g)(x)| \leq |f_n(x) - f(x)| + |g_n(x) - g(x)| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

So, $(f_n + g_n)$ converges uniformly to $f + g$ on E.

If in addition (f_n) and (g_n) are sequences of bounded functions, then $(f_n g_n)$ converges uniformly on E.

To start, note that if every f_n and g_n is bounded, then we must have that f and g are bounded.

Since (f_n) converges uniformly to f on E, for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $\forall x \in E, \ \ \forall n \geq N, \ \ |f_n(x) - f(x)| < \varepsilon$. Now fix $n \geq N$ and note that since $|f_n(x)|$ is bounded (say by M), we know that $|f(x)|$ must be bounded by $M + \varepsilon$.

The same reasoning also says that g is bounded.

Now set M to be greater than $|f(x)|$ and $|g(x)|$ for all $x \in E$. Then, for any $\varepsilon > 0$, pick N_1 and N_2 such that:

- $\forall x \in E, \ \ \forall n \geq N_1, \ \ |f_n(x) - f(x)| < \varepsilon$
- $\forall x \in E, \ \ \forall n \geq N_2, \ \ |g_n(x) - g(x)| < \varepsilon$

Then for any $x \in E$,

$$\begin{aligned} |f_n(x)g_n(x) - f(x)g(x)| &\leq |f_n(x)g_n(x) - f_n(x)g(x)| + |f_n(x)g(x) - f(x)g(x)| \\ &\leq |f_n(x)|\varepsilon + |g(x)|\varepsilon \leq (M+\varepsilon)\varepsilon + M\varepsilon = 2M\varepsilon + \varepsilon^2 \end{aligned}$$

Now ε and x are arbitrary. So, we conclude that $f_n g_n$ converges uniformly to fg on E.

Exercise 7.3: There exist sequences of functions (f_n) and (g_n) which converge uniformly on E such that $(f_n g_n)$ converges pointwise but not uniformly on E.

Set $E = (-1, 1) \setminus \{0\}$ and define $f_n = \frac{1}{|x|}$ for all $x \in E$ and $n \in \mathbb{N}$. Then trivially, (f_n) converges uniformly to $\frac{1}{|x|}$ on E. Also, define $g_n(x) = \frac{1}{n}$ for all $n \in \mathbb{N}$ and $x \in E$. Then clearly, g_n converges uniformly to 0 on E.

Claim 1: $(f_n g_n)$ converges pointwise to 0.

Fix any $x \in E$ and let $\varepsilon > 0$. Next, pick $N \geq \frac{1}{\varepsilon|x|}$. Then, for all $n \geq N$, we have that $|f_n(x)g_n(x) - 0| = |\frac{1}{n|x|} - 0| < \varepsilon$.

Claim 2: $(f_n g_n)$ does not converge uniformly to 0.

For any N, pick $0 < x < \frac{1}{N}$. Then $f_N(x)g_N(x) = \frac{1}{Nx} > 1$.

Exercise 7.4: Consider $f(x) = \sum\limits_{n=1}^{\infty} \frac{1}{1+n^2x}$.

- $f(x)$ converges absolutely for all $x \notin \{\frac{-1}{n^2} \mid n \in \mathbb{N}\} \cup \{0\}$.

  $f(x)$ is not defined for $x \in \{\frac{-1}{n^2} \mid n \in \mathbb{N}\}$. Also, $f(0) = \sum\limits_{n=1}^{\infty} 1$ doesn't converge.

  When $x > 0$, we immediately have that $f(x)$ converges (absolutely) by comparison with $\sum \frac{1}{n^2x}$.

  When $x < 0$, consider picking $N \in \mathbb{N}$ such that $N > \max\left(\frac{1}{\sqrt{-x}}, \sqrt{\frac{2}{-x}}\right)$.
  Then $f(x) = \sum\limits_{i=1}^{N} \frac{1}{1+n^2x} + \sum\limits_{i=N+1}^{\infty} \frac{1}{1+n^2x}$ where the left sum is finite and all the terms in the right sum are negative. Also in the right sum, $\left|\frac{1}{1+n^2}\right| = \frac{1}{n^2} \cdot \frac{1}{-x-\frac{1}{n^2}}$ where $0 \leq \frac{1}{-x-\frac{1}{n^2}} \leq \frac{1}{-\frac{x}{2}} = \frac{2}{-x}$. So by theorem 3.42 (proposition 57 in 140A notes), the right sum converges absolutely.

- f converges uniformly on $[r, \infty)$ and $(-\infty, -r] \setminus \{\frac{-1}{n^2} \mid n \in \mathbb{N}\}$ for all $r > 0$.

  Let $r > 0$ such that $r \neq \frac{1}{n^2}$ for any $n \in \mathbb{N}$.
  First, we'll show that f converges uniformly on $[r, \infty)$.

  Note that for all $x \in [r, \infty)$, $\left|\frac{1}{1+n^2x}\right| \leq \frac{1}{n^2r}$ and $\sum \frac{1}{n^2r} < \infty$. So by the Weierstrass M-test, we know that f converges uniformly on $[r, \infty)$.

  Next, we'll show that f converges uniformly on $(-\infty, -r] \setminus \{\frac{-1}{n^2} \mid n \in \mathbb{N}\}$.

  For $n \geq \max\left(\frac{1}{\sqrt{r}}, \sqrt{\frac{2}{r}}\right)$, we have that $1 + n^2x \geq 1 + \frac{1}{r}x < 1 - \frac{1}{r}r = 0$.
  So, $\left|\frac{1}{1+n^2x}\right| = \frac{1}{n^2} \cdot \frac{1}{|x+\frac{1}{n^2}|} = \frac{1}{n^2} \cdot \frac{1}{-x-\frac{1}{n^2}}$.

  Then $\frac{1}{n^2} \cdot \frac{1}{-x-\frac{1}{n^2}} \leq \frac{1}{n^2} \frac{1}{r-\frac{1}{n^2}} \leq \frac{1}{n^2} \cdot \frac{1}{r-\frac{r}{2}} = \frac{1}{n^2} \cdot \frac{2}{r}$ and $\sum \frac{1}{n^2} \cdot \frac{2}{r} < \infty$.

  So by the Weierstrass M-test, f converges uniformly on $(\infty, r]$.

This shows that f converges uniformly on any set which excludes all $x \in \{\frac{-1}{n^2} \mid n \in \mathbb{N}\}$ and whose closure doesn't contain 0.

- f does not converge uniformly on $(0,r]$ or $[-r,0) \setminus \{\frac{-1}{n^2} \mid n \in \mathbb{N}\}$ for any $r > 0$.

  Firstly, we'll show this for $(0,r]$:

  Note that for all $n \in \mathbb{N}$, at $x = \frac{1}{n^2}$ we have that $\frac{1}{1+n^2(\frac{1}{n^2})} = \frac{1}{2}$. Thus for all $n \in \mathbb{N}$, we can pick x such that $\left|\sum_{i=1}^{n} \frac{1}{1+xn^2} - \sum_{i=1}^{n-1} \frac{1}{1+xn^2}\right| > \frac{1}{3}$. In other words, the Cauchy Criterion is not fulfilled.

  Note, this also show that for any $N \in \mathbb{N}$:

  $$f(\tfrac{1}{n^2}) = \sum_{n=1}^{\infty} \frac{1}{1+\frac{n^2}{N^2}} \geq \sum_{n=1}^{N} \frac{1}{1+\frac{n^2}{N^2}} \geq \frac{N}{2}$$

  So f is not bounded on $(0,r]$.

  Similarly for $[-r,0) \setminus \{\frac{-1}{n^2} \mid n \in \mathbb{N}\}$:

  For all $n \in \mathbb{N}$, at $x = \frac{1}{-2n^2}$ we have that $\frac{1}{1-n^2(\frac{1}{2n^2})} = 2$. Thus for all $n \in \mathbb{N}$, we can pick x such that $\left|\sum_{i=1}^{n} \frac{1}{1+xn^2} - \sum_{i=1}^{n-1} \frac{1}{1+xn^2}\right| > 1$. So, the Cauchy Criterion is not fulfilled.

- Finally, all of this tells us that f is continuous.

  For any $x \in \mathrm{dom}(f)$, we know that f converges uniformly on a neighborhood around x. And since all the partial sums of f are continuous, we thus have that f is continuous on that neighborhood. So, f is continuous at x even when considering the whole domain of f.

Well, I ran out of time. Have fun not having to grade 4 additional questions. <3