Math 140B Lecture Notes (Professor: Brandon Seward)

Isabelle Mills

April 21, 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 $\mathrm{dom}(f)$.

Fix a point $x \in E \cap E'$. Then consider the function $\frac{f(t)-f(x)}{t-x}$ for $t \in \mathrm{dom}(f) \setminus \{x\}$ and define the derivative of f at x to be $f'(x) = \lim\limits_{t \to x} \left(\frac{f(t)-f(x)}{t-x}\right)$ 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 $f'(x) = \lim\limits_{t \to x} \left(\frac{f(t)-f(x)}{t-x}\right)$ is called the derivative of f.

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

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

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

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) $$(f+g)'(x) = f'(x) + g'(x) \qquad \text{(sum rule)}$$

(B) $$(fg)'(x) = f'(x)g(x) + f(x)g'(x) \qquad \text{(product rule)}$$

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

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:

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

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 $\text{dom}(f), \text{dom}(g) \subseteq \mathbb{R}$. 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 $(g \circ f)'(x) = g'(f(x))f'(x)$.

<u>Intuition</u>:

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

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 $f'(a) < \lambda < f'(b) \Longrightarrow g'(a) < 0 < g'(b)$. 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 $\mathrm{dom}(f) \subseteq \mathbb{R}$. If f' is defined and is itself differentiable, then the derivative of f' is denoted f'' and called the second derivative 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 $P = \{x_0, x_1, \ldots, x_n\}$, we write $\Delta x_i = x_i - x_{i-1}$ 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 $[a,b]$. Then, we define for each $i \in \{1, \ldots, n\}$:

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

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

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

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

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

Finally, we define:

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

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

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

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

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:

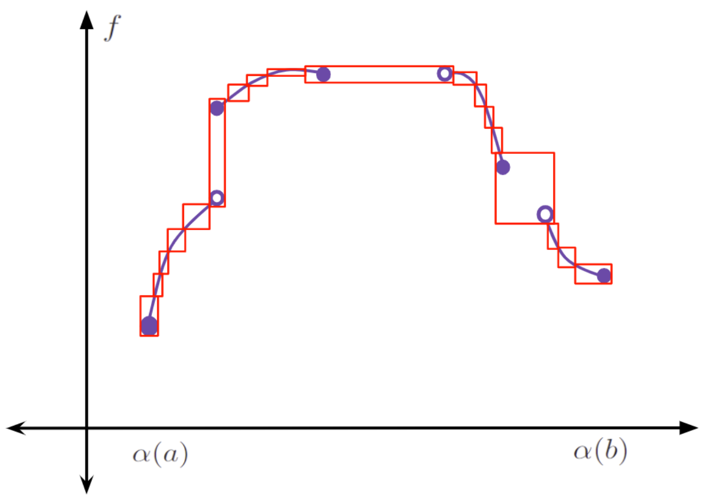
$$L(P,f,\alpha) \leq \sum_{i=1}^{n} f(s_i)\Delta\alpha_i \leq U(P,f,\alpha).$$

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 $\forall s,t \in K,\ \ |s-t| < \delta \implies |f(s) - f(t)| < \varepsilon$. So, pick any partition $P = \{x_0, \ldots, x_n\}$ of $[a,b]$ such that:

- $\{u_1, v_1, u_2, v_2, \ldots, u_m, v_m\} \subseteq P \subseteq K$
- $x_{i-1} \notin \{u_1, u_2, \ldots, u_m\} \implies \Delta x_i < \delta$ 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 $P=\{x_0,\dots,x_n\}$ of $[a,b]$ such that $U(P,f,\alpha)-L(P,f,\alpha)<\delta^2$. So, for each $i\in\{1,\dots,n\}$, 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 $A=\{i\in\{1,\dots n\}\mid M_i-m_i<\delta\}$ and $B=\{1,\dots,n\}\setminus A$.

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, $\sum\limits_{i\in B}\Delta\alpha_i<\delta$. 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 $\alpha(b) - \alpha(a) + 2 \sup_{m \le t \le M} |\phi(t)|$ is a constant.

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

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	
103		104	
105		106	
107		108	
109		110	
111		112	

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:

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

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:

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

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 $M = \sup\limits_{\gamma \leq x \leq x_0} |f(x)|$.

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 $\lim\limits_{t \to x} f'(t) = 0$. 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 $\lim\limits_{t \to x} f''(t) = 4$. 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 $\mathbb{R}^2$:
$$(x_1, x_2) \to (x_2, x_2) \to (x_2, x_3) \to \dots$$

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 $P = \{x_0, \dots, x_n\}$ of $[a, b]$ such that $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$.

Next, for each $i \in \{1, \dots, n\}$, 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.