LECTURE 22: THE MEAN VALUE THEOREM

1. Use the Chen Lu!

Video: Proof of the Chen Lu

Chain Rule (Chen Lu)

$$(g \circ f)'(x) = g'(f(x))f'(x)$$

Proof: From last time we have:

$$\frac{g(x+h) - g(x)}{h} = g'(x) + O(h)$$

Where $O(h)$ goes to 0 as $h \to 0$. This implies:

$$g(x+h) - g(x) = h\left(g'(x) + O(h)\right)$$

Use this with $f(x)$ instead of x and $f(x+h) - f(x)$ instead of h:

$$g(\cancel{f(x)} + f(x+h) - \cancel{f(x)}) - g(f(x)) = (f(x+h) - f(x))\left[g'(f(x)) + O(f(x+h) - f(x))\right]$$

$$g(f(x+h)) - g(f(x)) = (f(x+h) - f(x))\left[g'(f(x)) + O(f(x+h) - f(x))\right]$$

Note: This is valid *even if* $f(x+h) = f(x)$, which corrects the faulty proof from last time.

Date: Thursday, November 11, 2021.

Dividing both sides by h, we get:

$$\underbrace{\frac{g(f(x+h))-g(f(x))}{h}}_{\to (g\circ f)'(x)} = \underbrace{\left(\frac{f(x+h)-f(x)}{h}\right)}_{\to f'(x)} \left(g'(f(x)) + \underbrace{O(\underbrace{f(x+h)-f(x)}_{\to 0})}_{\to 0} \right)$$

Now if $h \to 0$, we get $f(x+h) - f(x) \to 0$ (by continuity of f) and therefore $O(f(x+h)-f(x)) \to 0$ (by definition of O), hence we obtain

$$(g\circ f)'(x) = f'(x)g'(f(x)) = g'(f(x))f'(x) \quad \square$$

For the rest of today, we'll prove a couple of theorems related to derivatives, such as Rolle's Theorem and the Mean Value Theorem.

2. Fermat's Theorem

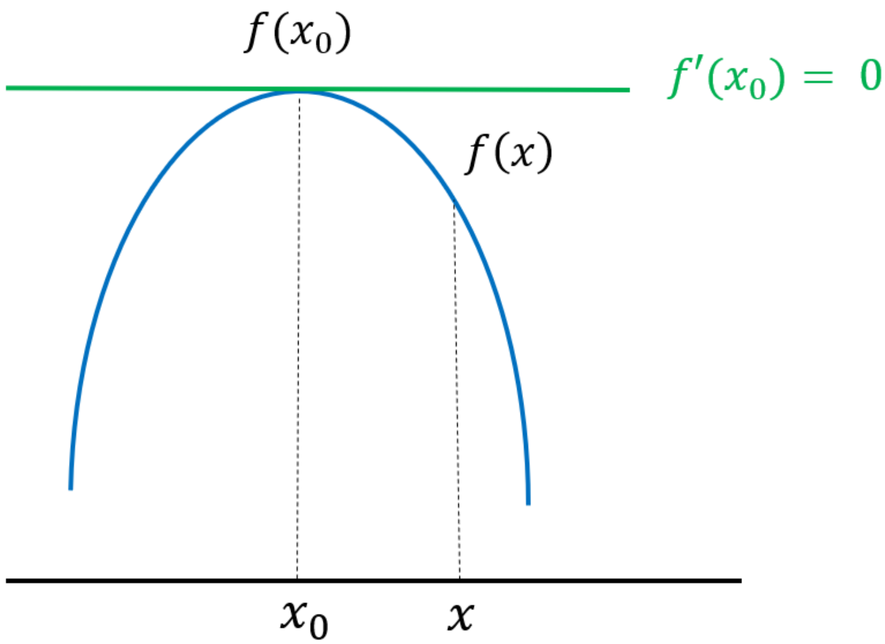
Video: Rolle's Theorem

Definition

f has a **local max** at x_0 if $f(x) \leq f(x_0)$ for all x near x_0
(similar for local min and strict local max/min)

Fermat's Theorem

If f is differentiable on (a,b) and has a local max or min at x_0, then $f'(x_0) = 0$



It's *this* theorem that makes optimization problems possible! It's because of this that you have to find the critical points of f, that is, points where $f'(x_0) = 0$ or where $f'(x_0)$ is undefined.

Proof: Assume WLOG that f has a local max at x_0 (replace f with $-f$ otherwise).

Then, by the definition of a derivative:

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to (x_0)^+} \underbrace{\frac{f(x) - f(x_0)}{x - x_0}}_{\leq 0} \leq 0$$

Here we used $f(x) \leq f(x_0)$ for x sufficiently close to x_0 since f has a local max, and so $\frac{f(x)-f(x_0)}{x-x_0} \leq 0$ (since $x > x_0$ by assumption).

Therefore $f'(x_0) \leq 0$, being the limit of a negative function.

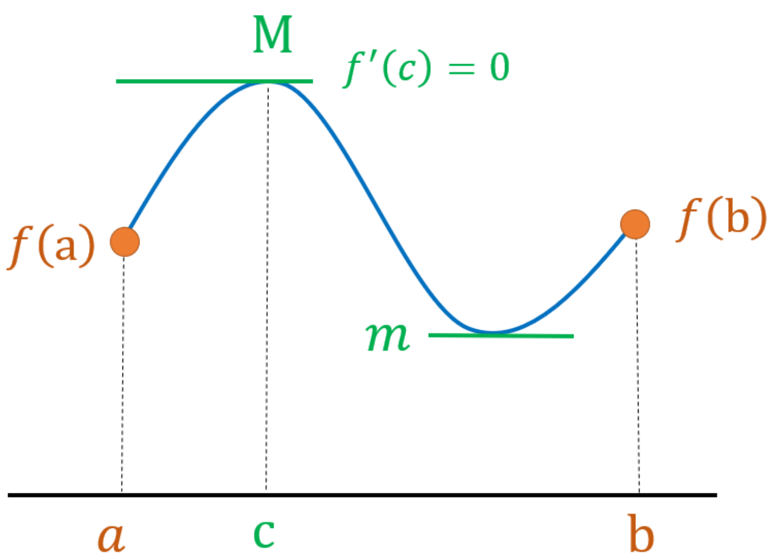
Similarly, considering the limit as $x \to (x_0)^-$ we get $f(x_0) \geq 0$, and so $f'(x_0) = 0$ $\square$

3. Rolle's Theorem

The next theorem will have you Rolle on the floor laughing ☺. It can be viewed as a special case of the Mean Value Theorem:

Rolle's Theorem

Suppose f is continuous on $[a, b]$ and differentiable on (a, b). If $f(a) = f(b)$, then there is some c in (a, b) with $f'(c) = 0$



Proof: Easy! Since f is continuous on $[a, b]$, by the Extreme Value Theorem, f must have a max M and a min m on $[a, b]$. We cannot have M and m be both a and b, otherwise $m = f(a) = f(b) = M$, and

f would be constant. Therefore f must have a max or a min at some point c in (a, b) and by Fermat's Theorem, we have $f'(c) = 0$ □

4. Mean Value Theorem

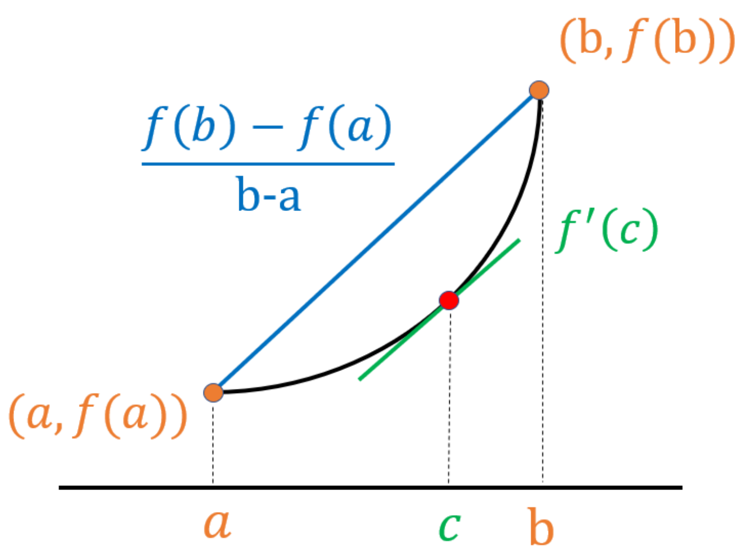
Video: Mean Value Theorem

We are now ready to state the third and final Value Theorem: The Mean Value Theorem. It can be viewed as the bigger sibling of Rolle, but surprisingly we can use Rolle to prove the MVT!

Mean Value Theorem

If f is continuous on $[a, b]$ and differentiable on (a, b). Then there is c in (a, b) such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$



Interpretation: MVT says that there is some point on your car trip where your instantaneous velocity $f'(c)$ equals to your average velocity $\frac{f(b)-f(a)}{b-a}$

Note: If $f(b) = f(a)$, then $\frac{f(b)-f(a)}{b-a} = 0$ so $f'(c) = 0$ and we recover Rolle's Theorem

Proof: The idea is to apply Rolle's theorem to a special function. Notice the equation of the line connecting $(a, f(a))$ and $(b, f(b))$ is

$$\text{Secant } = f(a) + \left(\frac{f(b)-f(a)}{b-a}\right)(x-a)$$

$$\text{Let: } g(x) = f(x) - \text{ Secant } = f(x) - f(a) - \left(\frac{f(b)-f(a)}{b-a}\right)(x-a)$$

$$\text{Then: } g(a) = f(a) - f(a) - \left(\frac{f(b)-f(a)}{b-a}\right)(a-a) = 0$$

$$g(b) = f(b) - f(a) - \frac{f(b)-f(a)}{b-a}(b-a) = f(b) - f(a) - f(b) + f(a) = 0$$

$g(a) = g(b)$ and by Rolle's Theorem, there is c with $g'(c) = 0$, that is

$$f'(c) - \left(\frac{f(b)-f(a)}{b-a}\right) = 0$$
$$f'(c) = \left(\frac{f(b)-f(a)}{b-a}\right) \quad \square$$

Let's illustrate the power of the MVT by showing 3 applications:

5. Application 1: Antiderivatives

Corollary

If $f'(x) = 0$ for all x in (a, b), then f is constant

Proof: Suppose not, then there are $x \neq y$ such that $f(x) \neq f(y)$, but then by the MVT, there is c between x and y such that

$$\begin{aligned}\frac{f(y) - f(x)}{y - x} =& f'(c) = 0 \\ f(y) - f(x) =& 0 \\ f(x) =& f(y)\end{aligned}$$

But this contradicts $f(x) \neq f(y) \Rightarrow\Leftarrow$ $\square$

Coro-Corollary

If $f'(x) = g'(x)$ for all x in (a, b), then $f(x) = g(x) + C$ for some constant C

Proof: Let $h(x) = f(x) - g(x)$, then $h'(x) = f'(x) - g'(x) = 0$, so $h(x) = C$, that is $f(x) - g(x) = C$ so $f(x) = g(x) + C$ $\square$

This says that two antiderivatives differ by a constant! This is why formulas like $\int x^2 dx = \frac{1}{3}x^3 + C$ are valid

6. Application 2: Increasing/Decreasing

Definition

f is **strictly** increasing if $x < y \Rightarrow f(x) < f(y)$
(similar for strictly decreasing, and increasing and decreasing)

Corollary

If $f'(x) > 0$ for all x, then f is strictly decreasing

Proof: Suppose $x < y$, then by the MVT, we get

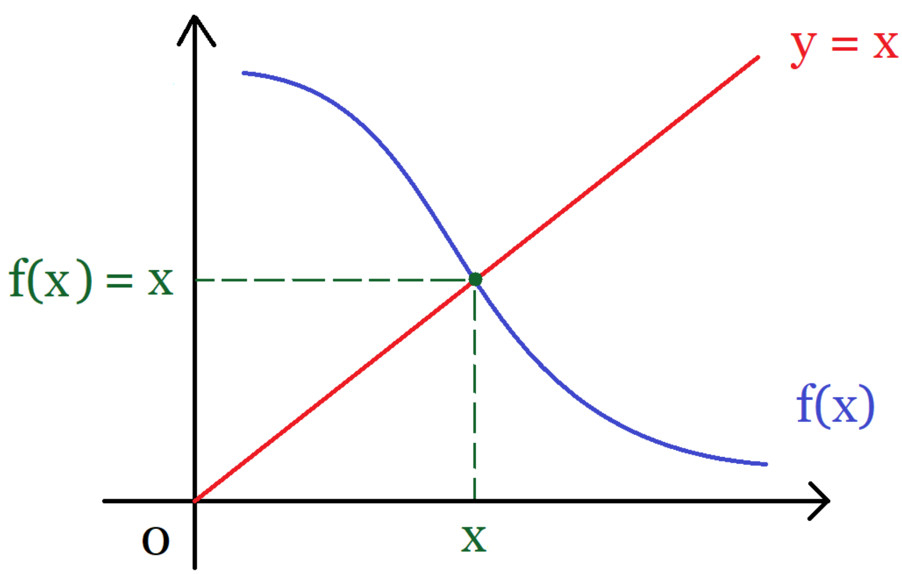
$$\frac{f(y) - f(x)}{y - x} = f'(c) > 0 \Rightarrow f(y) - f(x) > 0 \Rightarrow f(y) > f(x) \quad \square$$

7. Application 3: Fixed Points

Video: MVT and Fixed Points

Definition

f has a fixed point if $f(x) = x$ for some x



Corollary

If $f'(x) \neq 1$ for all x, then f has at most one fixed point

Proof: Suppose f has two fixed points a and b. Then $f(a) = a$ and $f(b) = b$, then by the MVT, we get

$$\begin{aligned}\frac{f(b)-f(a)}{b-a} =& f'(c)\\ \frac{b-a}{b-a} =& f'(c)\\ 1 =& f'(c)\end{aligned}$$

Which contradicts the fact that $f'(x) \neq 1$ for all $x \Rightarrow\Leftarrow$ □

Combining this with the fact from the section on the Intermediate Value Theorem, we get:

Theorem:

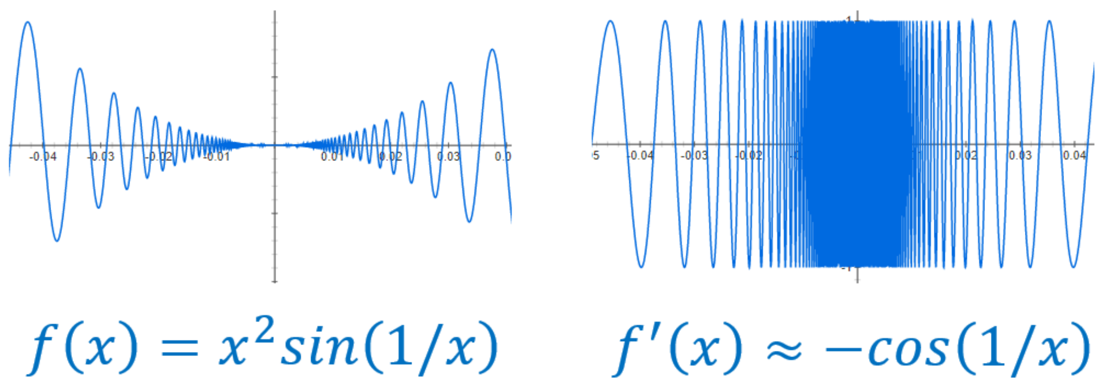
If $f : [0,1] \to [0,1]$ is continuous on $[0,1]$ and $f'(x) \neq 1$ for all x, then f has exactly one fixed point

8. Intermediate Value Theorem for Derivatives

Warning: If f is differentiable, f' doesn't have to be continuous!

Non-Example 1:

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$



Then $f'(0) = 0$ because:

$$f'(0) = \lim_{x\to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x\to 0} \frac{x^2 \sin\left(\frac{1}{x}\right) - 0}{x} = \lim_{x\to 0} x \sin\left(\frac{1}{x}\right) = 0 \text{ (Squeeze Thm)}$$

But f' is not continuous at 0 because

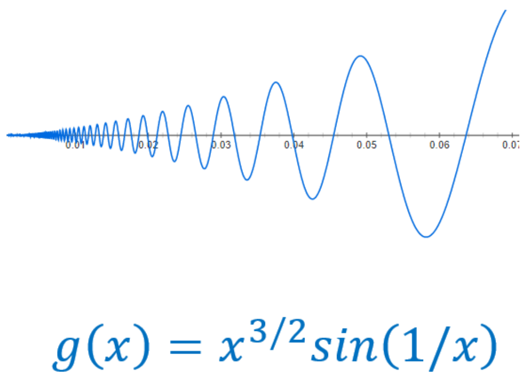
$$\lim_{x\to 0} f'(x) = \lim_{x\to 0} \underbrace{2x \sin\left(\frac{1}{x}\right)}_{\to 0} + x^2 \cos\left(\frac{1}{x}\right)\left(-\frac{1}{x^2}\right) = \lim_{x\to 0} -\cos\left(\frac{1}{x}\right) \text{ DNE}$$

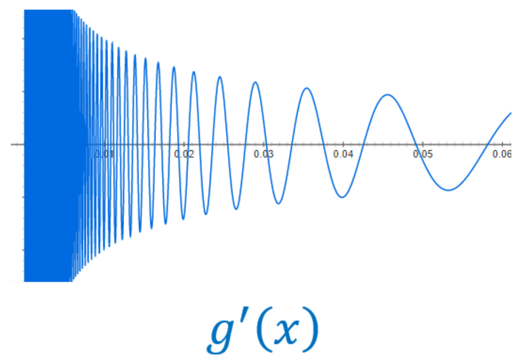
We certainly cannot have $\lim_{x\to 0} f'(x) = f'(0)$, and f' isn't continuous.

Even worse, can check that for the following example, $g'(x)$ blows up:

Non-Example 2:

$$g(x) = \begin{cases} x^{\frac{3}{2}} \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$



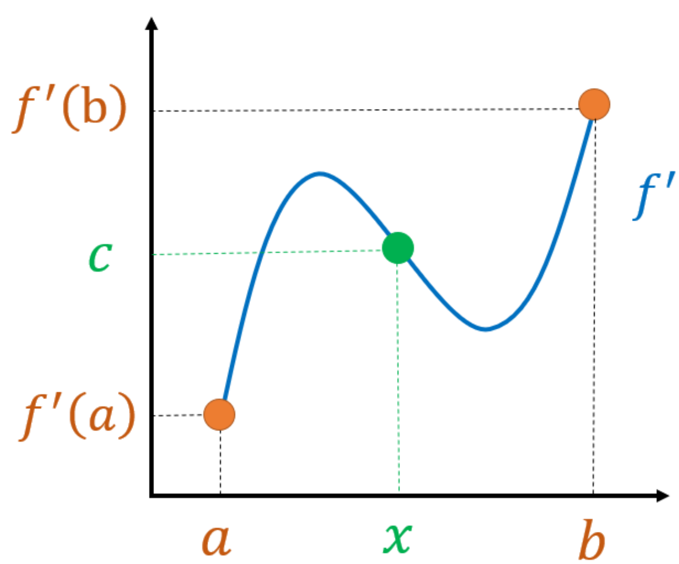


The surprising fact is that even if f' might be discontinuous, it still has the Intermediate Value Property, so it's not *that* bad after all:

IVT for Derivatives

If c is any number between $f'(a)$ and $f'(b)$, then there is x in (a, b) such that $f'(x) = c$

So f' can never have jump discontinuities (o/w it would omit values)



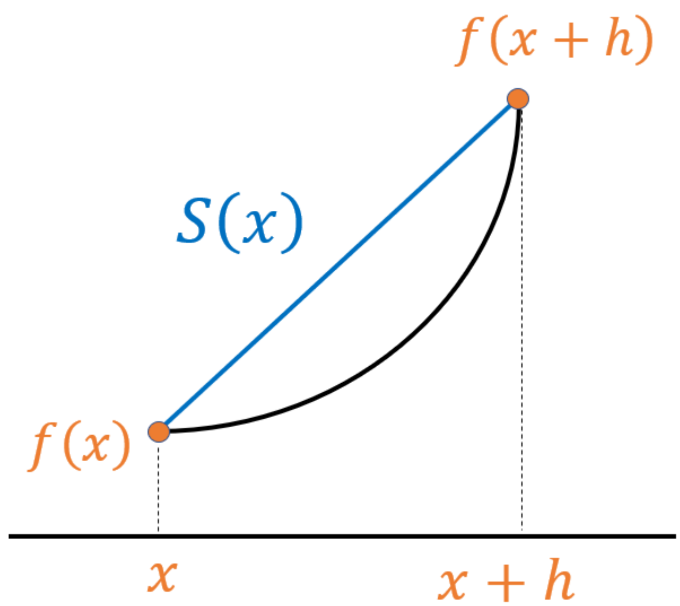
Proof: Beautiful proof that illustrates the "sliding secant method."

WLOG, assume $f'(a) < f'(b)$

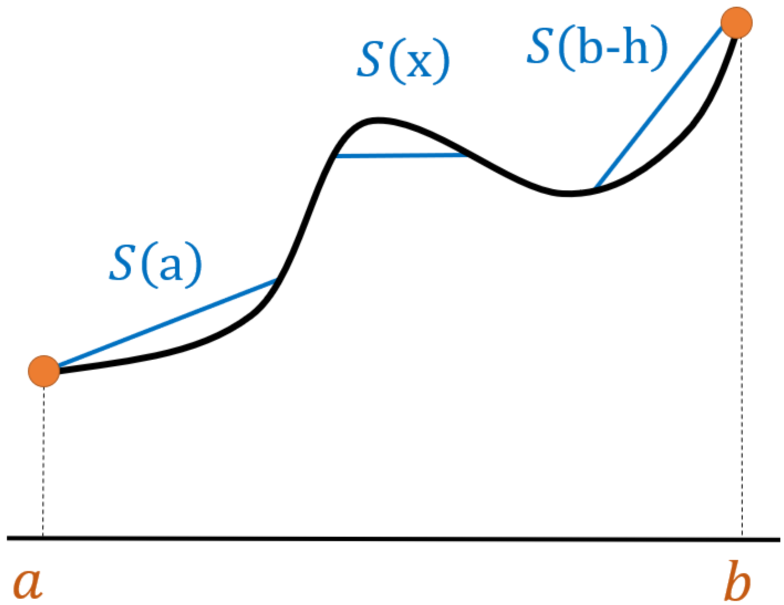
Let $h > 0$ be a fixed (but small) constant, and consider

$$S(x) = \frac{f(x+h) - f(x)}{h}$$

$S(x)$ is the slope of the secant line from $(x, f(x))$ to $(x+h, f(x+h))$.



The idea is to simply "slide" $S(x)$ from $x = a$ to $x = b - h$



From the Chen Lu Fact at the beginning, we have:

$$S(a) = \frac{f(a+h) - f(a)}{h} = f'(a) + O(h) \approx f'(a)$$

$$S(b-h) = \frac{f(b) - f(b-h)}{h} = f'(b) + O(h)\,^{1} \approx f'(b)$$

Since $f'(a) < c < f'(b)$, for h small enough we get $S(a) < c < S(b-h)$

But $S(x)$ is continuous, so by the IVT, there is x such that $S(x) = c$

$$\text{that is: } \frac{f(x+h) - f(x)}{h} = c$$

By the MVT, there is some p in $(x, x+h)$ such that $\frac{f(x+h)-f(x)}{h} = f'(p)$, so the above becomes

[1] possibly for a different function $O(h)$

$$f'(p) = c$$

Which is what we wanted (with p instead of x) $\square$