

[SQUEAKING]

[RUSTLING]

[CLICKING]

**CASEY** So let's continue with our discussion of the derivative. Let me recall that the derivative of a function, if it exists,  
**RODRIGUEZ:** which we denote by  $f'$  of  $c$ , this is the limit as  $x$  goes to  $c$  of  $f$  of  $x$  minus  $f$  of  $c$  over  $x$  minus  $c$ .

Last time, we looked at the relationship between continuity and differentiability and showed that differentiability implies continuity, but that having a derivative is quite a miraculous thing. Because there are continuous functions, which we actually constructed. We constructed a continuous function. You can generalize that example.

So we constructed a continuous function that is differentiable nowhere. So in fact, differentiability is a much stronger condition than continuity. And it's something of a miracle.

Just as a side note, if you go on to study complex analysis, which are functions of a complex variable rather than of a real variable, you also have the notion of a derivative there. And then, the derivative is much more miraculous than it is here in the setting of  $\mathbb{R}$ . But the whole point is, the derivative is something of a miracle.

So today, the goal is to prove some properties about the derivatives. Most of these you know. Maybe you didn't cover the proofs of them in calculus with the ultimate goal of proving the mean value theorem, which is, to me probably the most underappreciated but most important result in calculus.

Now, you could argue that probably the hero of the story of calculus is the fundamental theorem of calculus, which connects integration, which we'll cover next after differentiation. Which so the fundamental theorem of calculus connects integration and differentiation. So that really is, but-- and that is really the hero of calculus. It gets used more than anything else.

But to prove that, you use the mean value theorem. So if the fundamental theorem of calculus was Batman, I would say the mean value theorem is something like Alfred. He's really the reason why Batman gets to be who he is.

Enough analogies for now. Let's go back to proving some properties of the derivative. So first, let's do some basic-- the basic linearity and rules for the derivative. So let  $f$  go from  $I$  to  $\mathbb{R}$ .  $i$  is always some interval.

And if  $f$ -- so I have two functions,  $f$  and  $g$ , from  $I$  to  $\mathbb{R}$ .  $c$  is a point in  $i$ . And the conclusion is, if  $f$  and  $g$  are differentiable at  $c$ , then we have several rules that we can apply.

The first is basically the linearity of the derivative that for all  $\alpha$  and  $\mathbb{R}$ , function  $\alpha f$  plus  $g$ , which now goes from  $I$  to  $\mathbb{R}$ , is differentiable at  $c$ . And the derivative of the function  $\alpha f$  plus  $g$  prime  $c$  is  $\alpha$  times  $f$  prime of  $c$  plus  $g$  prime of  $c$ .

I think I'm starting to write a little slandered here. Second is the multiplication rule, which is the function  $f$  times  $g$  is differentiable at  $c$ . And the derivative of the product is not the product of the derivatives, but the prime of  $c$ ,  $g$  of  $c$  plus  $f$  of  $c$ ,  $g$  prime of  $c$ . So that's a multiplication rule.

And then, we also have the quotient rule. And just like when you need to divide by something, you need to assume that you're never dividing by 0. If  $g$  of  $x$  does not equal 0 for all  $x$  in  $I$ , then the function  $f$  over  $g$  is differentiable at  $c$ . And the derivative is the derivative of the top times the bottom minus the top times the derivative of the bottom over the bottom squared, [ $f$  now  $g$ ] prime  $g$  of  $c$  squared.

So I'll prove one and two. Three we'll leave as an exercise. So for one, we just compute that the limit as  $x$  goes to  $c$  of  $\alpha f$  plus  $g$  of  $x$ , minus  $\alpha f$  plus  $g$  of  $c$  over  $x$  minus  $c$ , this is just by definition  $\alpha$  times  $f$  of  $x$  plus  $g$  of  $x$  minus  $\alpha$  times  $f$  of  $c$  plus  $g$  of  $c$ . And so, this equals the limit as  $x$  goes to  $c$ .

And just collecting terms  $\alpha$  times  $\alpha f$  of  $x$  minus  $f$  of  $c$  times  $x$  minus  $c$ , plus  $g$  of  $x$  minus  $g$  of  $c$  over  $x$  minus  $c$ . And all of these limits exist. So this is just a fixed number. The limit as  $x$  goes to  $c$  of this exists and this exists.

So by what we know about limits, namely that the limit of the sum is the sum of the limits. And then, scalars just pull out of the limit. This is equal to  $\alpha$  times  $f$  prime of  $c$  plus  $g$  prime of  $c$ . So that's the proof of the first one.

For the proof of the second one, we'll use the fact that a function which is differentiable at a point is also continuous at a point. So since  $g$  is differentiable at  $c$ , it is continuous at  $c$ . I.e. limit as  $x$  goes to  $c$  of  $g$  of  $x$  equals  $g$  of  $c$ .

Now we compute the limit as  $x$  goes to  $c$  of  $f$  times  $g$  of  $x$  minus  $f$  of  $c$  times  $g$  of  $c$  over  $x$  minus  $c$ . Now what I'm going to do is add and subtract  $f$  of  $c$  times  $g$  of  $x$ . So this I can write as the limit as  $x$  goes to  $c$  of  $f$  of  $x$ , minus  $f$  of  $c$  over  $x$  minus  $c$ , times  $g$  of  $x$ . Plus now,  $f$  of  $c$  times  $g$  of  $x$  minus  $g$  of  $c$  over  $x$  minus  $c$ .

And again, so this limit as  $x$  goes to  $c$  exists. This is just the derivative of  $f$  at  $c$ . And we noted that since  $g$  is differentiable at  $c$ ,  $g$  is also continuous at  $c$ , meaning the limit as  $x$  goes to  $c$  of  $g$  of  $x$  equals  $g$  of  $c$ . This is just a constant  $f$  of  $c$ . This goes to  $g$  prime of  $c$ . So all these limits exist. And therefore, we get that this is equal to  $f$  prime of  $c$  times  $g$  of  $c$  plus  $f$  of  $c$ ,  $g$  prime of  $c$ .

And three, I'll leave as an exercise. Again, you'll write out-- you'll add and subtract something to  $f$  of  $x$  over  $g$  of  $x$  minus  $f$  of  $c$  over  $g$  of  $c$  and use the fact that differentiability at a point implies continuity, and just evaluate the limits. So I'm going to stop with these two rules.

I don't know why I called it the multiplication rule. It's called the product rule-- well, multiplication and product are the same thing. But we also have the chain rule, which requires a little more care to prove than what we've done so far for these rules.

So suppose I have two intervals. And then I have a function  $g$  going from  $I_1$  to  $I_2$ ,  $f$  going from  $I_2$  to  $R$ . And suppose  $g$  is differentiable at  $c$ , and  $f$  is differentiable at  $g$  of  $c$ . Then the function  $f$  of  $g$ , function goes from  $I_1$  to  $R$ . This is differentiable at  $c$ . And the derivative equals  $f$  prime of  $g$  of  $c$  times  $g$  prime of  $c$ .

So what's the basic idea? Again, so this basic idea is that we write the difference quotient,  $\frac{f(g(x)) - f(g(c))}{g(x) - g(c)}$  over  $x - c$  as  $\frac{f(g(x)) - f(g(c))}{g(x) - g(c)} \cdot \frac{g(x) - g(c)}{g(x) - g(c)}$ . And let  $x$  go to  $c$ , and then we'll pick up the derivative of  $g$  here at  $c$ .

And right here, this is  $f$  of something converging to  $g$  of  $c$ . And so, this should be like  $f'(g(c))$ . But the only problem with really writing this down is that there could be points where  $g(x) = g(c)$ . And then, I'm dividing by 0. That's not allowed. So writing this expression is meaningless.

But we're going to do something like this in spirit, where at least when  $g(x)$  does not equal  $g(c)$ , this thing kind of equals this thing. That's the basic idea.

So to implement this strategy, let's introduce some notation and a few auxiliary functions. So first off, the function we're interested in taking the derivative  $f$  of  $g$ , let me call that  $h$  of  $x$ . Because I don't want to have to keep writing  $f$  of  $g$  of  $x$ . Let's call it  $d$ . This will be the point  $g$  of  $c$ .

So we want to show then that  $h'(c)$  exists, and it's equal to  $f'(d) \cdot g'(c)$ . Just in this notation I've set up. So now, let me define some auxiliary functions  $u$  of  $y$ . This is going to be essentially the difference quotient of  $f$ , except at a certain point.

So this is  $\frac{f(y) - f(d)}{y - d}$  when  $y$  does not equal  $d$  so that this is meaningful. And then,  $f'(d)$  when  $y$  equals  $d$ . And then,  $v$  of  $x$ -- so this is going to play the role of that first expression I wrote up there,  $\frac{f(g(x)) - f(g(c))}{g(x) - g(c)}$ .

And then,  $v$  of  $x$ , this will be  $\frac{g(x) - g(c)}{x - c}$  when  $x$  does not equal  $c$ , and  $g'(c)$  when  $x$  equals  $c$ . Now, essentially what do we have here? Then if I look at  $f(y) - f(d)$ , this is equal to  $u(y) \cdot (y - d)$ , always.

Because if  $y$  does not equal  $d$ , then I divide over, and I get  $u$  of  $y$ , then I get  $f(y) - f(d)$  on the left-hand side over  $y - d$ . And that's equal to  $u$  of  $y$  by definition. And then, when  $y$  equals  $d$ , I get 0 here, and I get 0 over here. So that's clear. And also, the same here,  $g(x) - g(c) = v(x) \cdot (x - c)$ .

So we have these two expressions. And one more thing I want to note is that  $u$  and  $v$  are continuous at  $y = d$  and  $x = c$  respectively. So note that  $v$  of  $y$  is continuous at  $d$ , and  $v$  of  $x$  is continuous at  $c$ . So let's look at the proof of that. We just need to show that the limit as  $y$  goes to  $d$  of  $v$  of  $y$  equals  $v$  of  $d$ . So compute the limit as  $y$  goes to  $d$  of  $v$  of  $y$ .

Now, when we look at limits, remember, we're not allowed to put  $y = d$  in the expression for this limit. We're always looking at points  $y$  close to  $d$  but not equal to  $d$ , is what I'm saying. And so, therefore, when  $y$  is near  $d$ , and not equal to  $d$ ,  $v$  of  $y$  is given by-- I should-- this should-- I should have written  $u$  here, sorry,  $u$ .

So we're looking at  $y$  near  $d$ , but not equal to  $d$ . And therefore, it's given by this expression. So this is the limit as  $y$  goes to  $d$  of  $\frac{f(y) - f(d)}{y - d}$ . And this is just the definition of the derivative of  $f$  at  $d$ .

So remember,  $d$  is  $g(c)$ . So  $f$  is differentiable there. And this is by definition  $u$  of  $d$ . So this function  $u$  is continuous at  $y = d$ . And similarly, the limit as  $x$  goes to  $c$  of  $v$  of  $x$  equals  $v$  of  $c$ , which is  $g'(c)$ .

So now we're going to combine what I've written here to finish the proof. So in  $h$  of  $x$  minus  $h$  of  $c$ , this is equal to  $f$  of  $g$  of  $x$  minus  $f$  of  $g$  of  $c$ . And this is equal to  $u$  of  $g$  of  $x$  times  $g$  of  $x$  minus  $g$  of  $c$ . This is from this relation here. So we'll call this one star, this one double star. This is by star.

And now, I use the second relation, double star, to write  $g$  of  $x$  minus  $g$  of  $c$  equals  $v$  of  $x$  times  $x$  minus  $c$ . And therefore, the limit as  $x$  goes to  $c$  of  $h$  of  $x$  minus  $h$  of  $c$ , or  $x$  minus  $c$ , this is equal to the limit as  $x$  goes to  $c$  of  $u$  of  $g$  of  $x$  times  $v$  of  $x$ .

Now,  $u$  is continuous at  $g$  of  $c$ ,  $g$  is continuous at  $c$ . And we proved that the composition of two continuous functions is continuous. So this equals  $u$  of  $g$  of  $c$ . And  $v$  of  $x$ , we've proven already. That limit is  $v$  of  $c$ . And  $u$  of  $g$  of  $c$ , so this is  $d$ , remember, in the notation we had earlier. And therefore, that equals  $f$  prime of  $g$  of  $c$ . Remember,  $g$  of  $c$  was  $d$ , times  $g$  prime of  $c$ .

So to go from this point to this point, we use the fact that  $u$  is continuous at  $c$ , which we proved.  $u$  is continuous at  $g$  of  $c$ , which we proved over there. And that  $g$  is continuous at  $c$ . And therefore, the composition of two continuous functions is continuous at that point to arrive at this guy.

But again, the whole basic idea was to be able to write the difference quotient as a product of what looks like the difference quotient of  $f$  times the difference quotient of  $g$ . But you have to be careful, because  $g$  of  $x$  could equal  $g$  of  $c$ , in which case you're dividing by 0 in this basic idea we had up there, which is not allowed. But you get around it by using these functions which equal what you want as long as you're not dividing by 0, basically.

So that wraps it up for basic properties of the derivative. And now, we're going to move on to trying to prove the mean value theorem-- the Alfred of this Batman tale. But first, in order to prove this, we'll need a result about the derivative in relation to certain relative minimum and relative maximum.

So let me first define what this means. Let  $S$  be a subset of  $\mathbb{R}$ ,  $f$  be a function from  $S$  to  $\mathbb{R}$ . We say  $f$  has a relative max at  $c$  in  $S$ . If there exists a  $\delta$  positive, such that for all  $x$  in  $S$ ,  $x$  minus  $c$  is less than  $\delta$ , then  $f$  of  $x$  is less than or equal to  $f$  of  $c$ .

So the picture here, and then let me state that there's analogous definition of relative min at  $c$ . You just take this entire definition and flip the inequality. That's the definition for relative min.

So what this means pictorially, so remember, we had this notion of absolute min and absolute max. An absolute max at a point  $c$  means at every other point in  $S$ ,  $f$  of  $x$  lies below  $f$  of  $c$ . Here we're just saying, as long as you're close to the point  $c$ , then you lay below  $f$  of  $c$ .

So let's say our function looks something like that. So then, let's say we're at this point here. So this would be absolute max, because the graph of the function sits below the value of the function at this point for all  $x$  in this interval  $a, b$ . And then, here at this point, we would have an absolute min, because the rest of-- because the graph of the function sits above that point for all  $x$  in this interval.

But now, if I go to this point, say, this would be a relative max, because why? If I look at an interval-- small interval around it, and I look inside this interval-- there's too many dashed lines going on. So I'm going to erase this. Yeah, I'm just going to do away with the dash.

If I zoom in on just this one piece of the graph, then at this point, the graph sits below this point inside this little vertical strip. So this guy would be a relative max.

And then similarly, at this point here, maybe it's-- let me-- so you can really see what the graph is. I should have brought different colors. At this point, we have a relative min. Because nearby points sit above the value of the function evaluated at this point.

So relative max means nearby points  $f$  sits below that  $f$  evaluated at that point. And relative min is the other way around. So I don't think we need to belabor the point of what relative min relative maxes are. I think that's pretty clear.

So the theorem is the following. If  $f$  from  $a, b$  has a relative min or max at a point  $c$  strictly inside this closed interval, in the open interval  $a, b$ , then and  $f$  is differentiable at  $c$ , then  $f'$  of  $c$  equals 0.

And this is kind of clear from this picture, although this might look more like a sawtooth, but it shouldn't, that the tangent at this point is horizontal. And the tangent at this point is horizontal, which just is the expression that we're just expressing that  $f'$  of  $c$  equals 0. Because  $f'$  of  $c$  is supposed to be the slope of the tangent.

So the idea of the proof is-- and we'll do it for relative max. Relative min you get by taking minus  $f$ . So we'll do it for suppose  $f$  has a relative max at  $c$  in  $a, b$ . Then so I'm just going to draw a picture to illustrate why this is the case, so or why what I'm about to say is the case. So we have  $a, b$ , we have  $c$ .

Then there exist  $\delta$  positive, say  $\delta_1$ , such that two things are true.  $x$  or  $c - \delta_1 < c < c + \delta_1$  is contained in  $a, b$ . Two, for all  $x$  in  $c - \delta_1, c + \delta_1$ ,  $f$  of  $x$  is less than or equal to  $f$  of  $c$ .

So  $\delta_1$  is chosen as the minimum of-- so maybe let me not call it  $\delta_1$ , let me just call it  $\delta$ . So first off, you know that  $a, b$ , open interval, this is an open set which we've encountered in the assignments. So you can find a  $\delta_2$  so that  $\delta_2 > 0$ , say, so that this interval centered at  $c$  is contained in  $a, b$ .

Now, since  $f$  has a relative maximum at this point, there exist another  $\delta_3$ , say, so that if I'm inside that interval, then  $f$  of  $x$  is less than or equal to  $f$  of  $c$  for all  $x$  in that interval. And I choose  $\delta$  to be the minimum of these two  $\delta$ s.

And then I will have a chosen  $\delta$  so that  $c - \delta < c < c + \delta$  is-- so here, this one is-- and then  $\delta_2$  shows them so that that is contained. So I choose the minimum of these  $\delta$ s. And this is the  $\delta$  I take for this statement.

So now, the way we get that the derivative is 0 is, we're going to approach  $c$  from above and from below and use the fact that we have a relative max to show that the derivative is both bigger than or equal to 0 and less than or equal to 0.

So let  $x_n$  be  $c - \frac{\delta}{2n}$ , which is in the interval  $c - \delta < c < c + \delta$  for all  $n$ . Then what do we know?  $x_n$ , the limit as  $n$  goes to infinity of  $c - \frac{\delta}{2n}$  converges to  $c$ . So the derivative at  $c$  is equal to the limit as  $n$  goes to infinity of  $\frac{f(x_n) - f(c)}{x_n - c}$ .

Now, what do we note? So  $x_n$  is in this interval here, which is contained in this interval here. And therefore,  $f$  of  $x_n$  is always less than or equal to  $f$  of  $c$ . So what's on top is always less than or equal to 0.

Now,  $x_n$  is always less than  $c$ . So this thing on bottom is also less than 0. And therefore, something that's less than or equal to 0 on top over something that's less than 0 on bottom must be bigger than or equal to 0. So that's approaching  $c$  from the left.

If we approach  $c$  from the right,  $\frac{\Delta}{2^n}$  which is contained in  $c, c + \Delta$  for all  $n$ , then  $x_n$  converges to  $c$ . And if we look at  $f'$  of  $c$ , again, this is equal to limit as  $n$  goes to infinity of  $\frac{f(y_n) - f(c)}{y_n - c}$ .

Again,  $f$  of  $c$  is a relative maximum. So this thing on top is still less than or equal to 0. And the thing on bottom though, is now since  $y_n$  is bigger than  $c$ , this is positive. So something on top is negative, something on bottom is positive. So the limit must be negative.

So we've just shown  $f'$  of  $c$  is both bigger than or equal to 0. And that it's also less than or equal to 0. And therefore, the derivative is 0 at this point.

And again, the proof for a relative minimum is very similar, except that inequality flips signs. And therefore, you would get this is less than or equal to 0, this one would be bigger than or equal to 0. But you still get  $f'$  of  $c$  is bounded between 0 and 0.

So at a point where we have a relative minimum strictly inside the interval, the derivative must be 0. And the fact that it happens strictly inside the interval is important. This is not necessarily true if the relative max or relative min happens at an endpoint.

Think of  $f(x) = x$  on  $[0, 1]$ . Then it has a absolute minimum at 0, absolute maximum at 1, and the derivative is 1 at both points, not 0. So this is only for relative min, relative maxes that occur inside-- strictly inside the interval that the function is defined on and where it's differentiable.

So now, we have Rolle's theorem, which is essentially the mean value theorem rotated. We'll get the mean value theorem from Rolle's theorem. And which states the following, let  $f$  be from  $a, b$  to  $\mathbb{R}$ , be continuous. So it's continuous at every point in  $a, b$ , differentiable at every point inside of  $a, b$ .

It could be differentiable at the endpoints. That's fine. But it has to be differentiable on the open interval  $a, b$ . If  $f(a) = f(b)$ , then there exists a point  $c$  in the open interval  $a, b$ -- and this is important-- such that  $f'(c) = 0$ .

And what's the picture that goes with this? I'm sure you've seen it before. We draw a function that looks kind of like sine or cosine. But anyways, function which is 0 here, 0 here. Then there has to exist a point where the tangent is vertical. In fact, here it can occur at two different points.

But something that's already kind of giving the game away as far as how we'll prove this is, let's look at where the function takes a maximum and a minimum. That is the heart of the proof. And why can we-- why does it even have an absolute maximum or absolute minimum?

So I guess I didn't say this over there when I discussed relative min and relative max. But an absolute maximum is also a relative maximum. And absolute minimum is also a relative minimum. So let me make this-- this is kind of a late remark. Absolute max is a relative max, absolute min is a relative min.

So this picture is kind of given the game away already. Let's look at where the function takes a max and a min. And why does  $f$  take a max or a min? It's because  $f$  is continuous on this closed and bounded interval. This is something we proved when we were discussing continuity, the min max theorem. So let's give the proof.

Since  $f$  is continuous on  $a, b$ ,  $f$  achieves a relative max at some point  $c_1$  in  $a, b$ , and relative min at  $c_2$  in  $a, b$ . If there exists an  $x$ -- also. If  $f$  of  $c_1$  is positive, then what can you say?

Well,  $c_1$  cannot be one of the endpoints, because at the endpoints,  $f$  of  $a$  equals  $f$  of  $b$  equals  $0$ . This implies  $c_1$  is in the open interval  $a, b$ , which implies by the previous theorem,  $f$  prime of  $c_1$  equals  $0$ . So here you--  $c$  equals  $c_1$ .

If I'm at the minimum now, and this is less than  $0$ , then again, this implies that  $c_2$  is in the open interval  $a, b$ , which implies since  $f$  has a minimum at  $c_2$ ,  $f$  prime of  $c_2$  equals  $0$  by the previous theorem. And therefore, we could take  $c$  equals  $c_2$ .

If-- now, this is two cases--  $f$  of  $c_1$ , if either  $f$  of  $c_1$  is positive or  $f$  of  $c_2$  is less than  $0$ , then we have the result. If  $f$  of  $c_1$  is less than or equal to  $0$ , is less than or equal to  $f$  of  $c_2$ , now remember,  $f$  achieves a max that  $c_1$ .  $f$  achieves a min at  $c_2$ . And therefore, this should always be less than or equal to this.

If the max and the min-- so first off, I want to make sure I'm getting my logic correct here. So rather than state it this way, let's do it like this.

So in terms of the picture here, this would be  $c_1, c_2$ , where we achieve a max and a min. Then in the final case that  $f$  of  $c_1$  and then, so this is the max.  $0$  is equal to  $f$  evaluated at a point. Then-- all right. So let me go back to what I was going to say a minute ago.

So in the last case, that  $f$  of  $c_1$  is less than or equal to  $0$  is less than  $f$  of  $c_2$ . Now, remember, this is the min, this is the max of  $f$ . So this should always sit above  $f$  of  $c_2$ . So that implies that  $f$  of  $c_1$  equals  $f$  of  $c_2$ . And therefore, the max and min equal each other.

Which implies for all  $x$  and  $a, b$ , since  $f$  of  $c_1$  is the minimum, or  $f$  of  $c_2$  is the-- so I started off with something here, there, and in between, which implies for all  $x$  in  $a, b$ ,  $f$  of  $x$  equals  $f$  of  $c_2$ . So that means  $f$  is constant. And we know what the derivative of a constant function is. It's just  $0$ .

So we could take-- say, the midpoint. And that's the end. Sorry that I kind of fumbled through that for a minute, but-- OK. So for any function which is  $0$  at the endpoints, there has to be a point in between where the function is-- the derivative of the function equals  $0$ .

Now, again, we should come back to when we see a theorem, we should pick it apart a little bit to see what's necessary, what's not. We had these two hypotheses coming in, that if the function is continuous and differentiable on this open interval  $a, b$ , are these necessary?

So for example, let's say I have the function-- if I look at the function  $1$  minus the absolute value of  $x$ . So at  $-1, 1, 1$ , this is a continuous function on  $[-1, 1]$ . And  $f$  of  $-1$  equals  $f$  of  $1$  equals  $0$ . But there is no point where  $f$  prime of  $c$  equals  $0$ .

What hypothesis did I leave out for this function? The fact that it's-- the hypothesis that it's differentiable at every point in between. This function is not differentiable. But there's no point where the derivative equals  $0$ .

So this example tells you that the hypothesis that the function is differentiable on the interval minus 1, 1, or a, b, is necessary for the theorem to be true. That's what I'm getting at.

Now, the other hypothesis that the function is continuous on the closed interval is also necessary. Let's say I look at-- and this is pretty easy to come up with a counter-example, if I don't assume that.

Let's take the function which is-- the function that's  $x - 1$  for  $x$  not equal to minus 1, and 0 for  $x$  equal to minus 1. So does this function look like?

So this function here is differentiable on minus 1, 1,  $f$  of minus 1 equals  $f$  of 1 equals 0. But again, there's no point  $c$  in minus 1, 1 so that  $f'$  of  $c$  equals 0. Because the derivative in between minus 1, 1 is just 1.

So both of these hypotheses, that  $f$  is continuous on the closed interval and differentiable on the open interval, are necessary for this theorem to be true. If you drop either of those hypotheses, then the theorem is false.

So let's rotate this picture here and arrive at the mean value theorem. Let  $f$  from  $a, b$  to  $\mathbb{R}$  be continuous. So it's continuous on the closed interval  $a, b$ , differentiable on  $a, b$ . Then there exists a  $c$  in the open interval  $a, b$  such that  $f'$  of  $c$ -- let me write it this way--  $f$  of  $b$  minus  $f$  of  $a$  times  $b$  minus  $a$ .

So why do I refer to it as a rotated picture and shifted version of Rolle's theorem? So there's  $a, b$ . Here's  $f$  of  $a, f$  of  $b$ . And let's say that's how the function looks. Now, the slope of the line connecting  $f$  of  $b$  to  $f$  of  $a$  is exactly  $f$  of  $b$  minus  $f$  of  $a$ .

So the slope of that line is  $f$  of  $b$  minus  $f$  of  $a$ . And what we're stating is that there's a point  $c$ , so that the tangent at that point is parallel to that line, has the same slope as the line connecting  $f$  of  $b$  to  $f$  of  $a$ .

So this theorem really does reduce to Rolle's theorem. So let me define a function  $g$  from  $a, b$  to  $\mathbb{R}$ , which satisfies the hypotheses of Rolle's theorem and will give us what we want, basically. So  $g$  of  $x$  equals  $f$  of  $x$  minus  $f$  of  $b$ , plus  $f$  of  $b$  minus  $f$  of  $a$  over  $b$  minus  $a$  times  $b$  minus  $x$ .

Basically, what this function  $g$  does is, it takes the function now and rotates it, and shifts it down, so that  $f$ -- these two points then coincide and give you 0. Then  $g$  is continuous on the closed interval  $a, b$ . Why? Because it's a sum of two continuous functions on  $a, b$ .  $f$  of  $x$  is continuous on  $a, b$ .

And this is just a constant. That's just a constant. This is a polynomial. So  $g$  is continuous on  $a, b$ , and differentiable on  $a, b$ , again, because  $f$  is.  $f$  is differentiable on the open interval  $a, b$ . This part is differentiable everywhere so the sum of two differentiable functions is differential. So  $g$  of  $x$  is differentiable on  $a, b$ .

So continuous on  $a, b$  and differentiable on  $a, b$ . And since I said, we're going to use Rolle's theorem, let's compute  $g$  of  $a$ . This is equal to  $f$  of  $a$  minus  $f$  of  $b$ , plus  $f$  of  $b$  minus  $f$  of  $a$ , over  $b$  minus  $a$ , times  $b$  minus  $a$ . That cancels with that. Minus  $f$  of  $b$  cancels with  $f$  of  $b$ .  $f$  of  $a$  cancels with minus  $f$  of  $a$ , and we get 0.

And then,  $g$  of  $b$ , this is even easier to see. If I just stick this in, this is  $f$  of  $b$  minus  $f$  of  $b$  plus  $f$  of  $b$  minus  $f$  of  $a$ , over  $b$  minus  $a$ ,  $b$  minus  $b$ . And this also equals 0. So  $g$  of  $a$  and  $g$  of  $b$  are 0. The function is continuous on the closed interval. It's differentiable on the open interval.



And therefore, by Rolle's theorem-- I think there's supposed to be some sort of accent over  $O$ , but I can't remember, and I've forgotten it already. By Rolle's theorem, there exist a point  $c$  in  $a, b$ , such that  $0$  is equal to  $g$  prime of  $c$ , which is equal to-- we take  $g$  and actually compute what the derivative is at  $c$ , is equal to  $f$  prime of  $c$ . That's a constant derivative of that is  $0$ .

This times this, the derivative of  $x$  evaluated at  $c$  just gives me minus  $f$  of  $b$  minus  $f$  of  $a$ , times  $b$  minus  $a$ . And that's it.  $0$  is equal to  $f$  prime of  $c$  minus what we want. We move that to the other side, multiply through by  $b$  minus  $a$ , and we're done.

So some first very nice applications of the mean value theorem, which I think you learned from calculus are the following. Let  $f$  from  $I$  to  $R$  be differentiable, meaning it's differentiable at every point in the interval  $I$ .

Then we have two conclusions.  $f$  is increasing, meaning we recall what that means, for  $f$  to be increasing, that means that if  $x$  is less than  $y$ , this implies  $f$  of  $x$  is less than or equal to  $f$  of  $y$ .

This is equivalent to for all  $x$  in  $I$ ,  $f$  prime of  $x$  is bigger than or equal to  $0$ . So remember,  $f$  prime is the rate of change. If the rate of change is always non-negative, then  $f$  has to be growing. But this says,  $f$  is decreasing, meaning  $x$  less than  $y$  implies  $f$  of  $x$  is bigger than or equal to  $f$  of  $y$ . This is equivalent to since a function is increasing if and only if minus  $f$  is decreasing, this inequality should flip.

So I'm going to do increasing-- decreasing follows from number one by taking minus  $f$ . So how much time do I have? I still have time. So suppose first that  $f$  prime of  $x$  is always non-negative. Let  $a, b$  be in  $I$ . [SNEEZES] Excuse me.

We now want to show  $f$  of  $a$  is less than or equal to  $f$  of  $b$ . Then  $f$  is continuous on this smaller interval  $a, b$ . Why? Because  $f$  is differentiable at every point in  $I$ . And therefore, it is continuous at every point in  $I$ . So in particular on this smaller interval, and differentiable on  $a, b$ , which implies there exists a  $c$  in  $a, b$ --

So by the mean value theorem, such that  $f$  of  $b$  minus  $f$  of  $a$  equals  $f$  prime of  $c$  times  $b$  minus  $a$ . And now, since the derivative is always non-negative and  $b$  minus  $a$  is positive, this is bigger than or equal to  $0$ . I.e. if  $a$   $b$  minus  $f$  of  $a$  is bigger than or equal to  $0$ .

So we prove in one direction. We've proven that if you like that the function's derivative being non-negative implies that the function is increasing.

So now let's prove the opposite direction. So suppose  $f$  is increasing and  $c$  is in  $I$ . Let  $x_n$  be a sequence in  $I$ , such that  $x_n$  converges to  $c$ . And one of two things holds. And either let's say,  $a$ , all  $n$   $x_n$  is less than  $c$ , or  $b$ , all  $n$   $x_n$  is bigger than  $c$ . And we can always find such a sequence given a point  $C$  in the interval  $I$ .

So if  $I$ , so let's say, it's a closed interval. And  $c$  is in here. Then we can find a sequence  $x_n$  converging to  $c$ , approaching  $c$  from the left. If  $c$  is one of the endpoints, then we can find a sequence from  $I$  approaching  $c$  from the right. That's  $b$ .

And the same thing if  $c$  is over here we can find a sequence approaching  $c$ , strictly from the left. So there always exists such a sequence. Always exists. And this is because,  $I$  is an interval.

And if the sequence satisfies case a, and we look at  $f'$ , then we get that  $f(x_n)$  is less than or equal to  $f(c)$  for all  $n$ . And let me in fact write this a little bit differently. Since-- so here we're using that  $f$  is increasing. Since  $f$  is increasing, which implies for all  $n$ ,  $f(x_n) - f(c) < 0$ . Now, this thing on top is negative.  $x_n$  is less than  $c$ .

So this thing on the bottom is also negative. And therefore, this thing is, positive-- non-negative, I mean. And therefore, the limit must also be negative. But that limit is just a derivative evaluated at  $c$ .

And case b is kind of similar. I mean, not kind of, it is. In case b, we get that for all  $n$ ,  $f(x_n)$ , now, since  $x$  tends to the right of  $c$ -- let me write it this way.  $f(x_n) - f(c) > 0$ , which implies that  $f'(c)$ , which is the limit as  $n$  goes to infinity of  $f(x_n) - f(c)$  over  $x_n - c$ . What's on top is non-negative.

And what's on bottom is non-negative, because  $x_n$  is bigger than  $c$ . So bigger than 0, bigger than 0, bigger than 0. Thus in either case, we get that  $f'(c)$  is bigger than or equal to 0. And therefore, we've proven that the derivative is always non-negative. So that's one.

And two,  $f$  is decreasing if and only if the function minus  $f$  is increasing, which by what we've done from part one is if and only if  $-f'(x)$  is bigger than or equal to 0 for all  $x$  in  $I$ . And then, multiplying through by this minus 1 gives me number two.

And so, let me make one last-- I know it's taboo to write on the back part. But I'm going to do it anyways. We have the very simple theorem which now follows. Let  $f$  from  $I$  to  $\mathbb{R}$  be differentiable, then  $f$  is constant. So  $f(x) = f(y)$  for all  $x$  and  $y$  in  $I$ , if and only if the derivative is identically 0.

So  $f$  is constant if and only if  $f$  is both increasing and decreasing. Because it satisfies the equality sign in both of those.

And by what we've done now, this means for all  $x$  in  $I$ ,  $f'(x) > 0$ , and-- that's for the increasing part--  $f'(x) < 0$ . And therefore, this is equivalent to  $f'(x) = 0$ . So we'll stop there.