[SQUEAKING]

[RUSTLING]

[CLICKING]

TOBIAS COLDING: All right. So last time, we were looking at a function from, say, R, or it could be some interval in r. And then we have a point x in R.

And we say that f is differentiable at x0 if when we form this different quotient-- so this quotient here-- if this here has a limit x go to x0.

And so, again, this means that there is some number so that if x is sufficiently close, is delta close to x0, then this quotient here is close to that number. And it's always assumed that x is not equal to so that the denominator makes sense. So that is what it means for the function to be differentiable.

And then one of the first things we saw was that we had a lemma that if f is differentiable at x0, then it implies that it's also continuous at x0.

And just very briefly, we had four rules that if you had two functions, f and g, and they are both differentiable and you're looking at where you're adding the two functions, then the sum here of the two functions is also differentiable. And the derivative is the sum of the derivative.

Another thing we had was that if you take a constant times a function and this function is differentiable, then this new function is also differentiable. And the derivative is that constant times the derivative of the original function.

And then the third thing was the Leibniz rule. And that was if you have these two functions that are differentiable and you take the product and you take the derivative, that exists and is equal to the derivative of the first times the second plus the first times the derivative of the second.

And then the fourth of these was the quotient rule. And the quotient rule was that you have, again, these two functions. You assume that g here is not equal to 0 at any point. And then you're looking at the quotient here. And the quotient is also differentiable. And the derivative is given by this formula.

This is the formula for the derivative, where this is the function multiplied by itself. And then there was a last rule that was the chain rule. So the chain rule-- and that was that if you take a function, say, from some interval into some set, let's call it-- let's say that the interval that the function is defined is some interval from a to b.

And this function takes value in another interval from c to d. And then you have a function g that is defined on c to d into R. Then you can look at the composition. You take some x, and you first apply f to it. And then afterwards, you apply g. And this is typically written as this.

And the chain rule was that if f and g are differentiable, then the composition here is also differentiable. And the derivative of the composition is the derivative of the second function evaluated at this point here times the derivative of the first in x. So this was the chain rule. OK. That was just a quick recap.

And now the first thing we'll be interested in is if we have a-- so suppose that f here is a function, let's say defined on some interval from a to b into R. And f is differentiable.

It doesn't need to be differentiable for this here to make sense. But we say that x0-- and x0 is now in the open interval here.

So we say that x0 is-- so we say that f has a local maximum at x0 if there exists delta greater than 0 such that if you restrict f and you just think about f on this little interval around x0-- so you're forgetting about how f looks like everywhere else.

Then if you're looking at the maximum of this function-- so this is the same as just the maximum of f on this little interval. Then the maximum is achieved at x0 if this is the case. So you have the picture here. You have a function. We talked about this a little bit last time. You have a function. Maybe the function looks like this.

And this point here is x0. And so this here is the value at x0. And you see that there's some little interval-- say this interval here. If you're looking at the function in this little interval, then it's clear that the maximum is at x0. But, of course, it's just a local maximum. If you were allowed to move further away over here, you see that the value of the function is much larger.

So this is a local maximum. And for this to make sense, you don't need the function to be differentiable. But when we apply it, we will assume that it is differentiable. So this is the local maximum. And this is the picture, that is it's just a maximum if you're looking at values near x0, but it's not a global maximum. Of course, it could be a global maximum.

And then similarly, for a local minimum-- and I'll just draw the picture because the definition is kind of obvious. It's the same, that you take a function here. And so the function maybe could look like this.

And you see that if this here was x0, then in some little neighborhood around x0, the value here would indeed be the minimum. But if you're moving much further away, then there might be other points that are even smaller, where the value is even smaller. So this is like a local minimum. And so the first observation you have about this is the following lemma.

If f here is a function, say, on an interval like this, it's differentiable and has a local maximum or local minimum at x0.

And now here, if I'm thinking about a function defined on a closed interval like this, it's important that where the local max or min is in an interior point. It's not allowed to be one of the boundary points.

So if it is differentiable and it has a local maximum or minimum at such an interior point, then the claim is that then the derivative at x0 is equal to 0. And so now let's try to prove that.

So we have here-- so the picture is-- let's just assume. So proof-- assume that f has a local maximum at x0. If it's a local minimum, the argument is the same with obvious modifications.

OK. So you have some picture, like this. Again, here, you have x0. And now this here is the function that-- oops, maybe this. So this is the function right. We see that at the value here, at x0, it's maximum. And then who cares? We don't care about what happen when you are further away.

And so you divide into two cases. You're looking at where x here is bigger than x0, if x is bigger than x0. And now you're looking at-- so then you have that f of x0 minus f of x0. The value here is maximum.

So this value here is-- so if this is the case and x is sufficiently close to x0-- if it's very close here, then it has a maximum at this point. So this means that this difference here is non-positive.

This didn't really use that it was strictly bigger than, right? It just used that it was close to x0. But now this is where we're using that it's bigger than-- because you also have that x minus x0-- and, again, this is where this assumption's used-- this thing here is positive.

So this means that if you take something nonpositive and you divide it by something positive, then what you get is something that is nonpositive. So this implies that for the difference quotient, f of x minus f of x0, x minus x zero-- again, this here is nonpositive. This is positive, which means that the whole thing is nonpositive.

So as you're going towards x0 from the right, the difference quotient here is nonpositive. If you now go to x0 from the left, if you go in this direction-- so this means that now you have x here is less than x0 and sufficiently close.

So if you have that, then well, you again have-- this was just because it was a local maximum. You still have that this thing here is not positive. But now we have that x minus x0, this thing here now is-- because of this, this means that this thing here is strictly less than 0.

So you take something that is nonpositive and you divide it by something negative, then you get something nonnegative. So therefore, this implies that for the difference quotient, f of x minus f of x0, for the difference quotient here, you get that this thing here, this is nonpositive. This is negative. This means that this here is nonnegative.

So as you're coming in-- so you see that as you're approaching this direction to x0, then the difference quotient is not positive. When you're coming in here, it's nonnegative. But it has a limit. And so the limit from this side would have to be-- the limit from this side would have to be nonpositive and the limit from this one nonnegative. So therefore, the limit must be 0.

So we conclude that the limit here-- hence, the limit here, as x goes to x0 of f of x of the difference quotient must be 0. And so this is the derivative at that point-- and similarly if it was a local minimum.

And at point where the derivative-- so maybe I'll just say that-- so if f prime at x0 is equal to 0, then we often say that then x0 is said to be a critical point for f.

And let me just give you an example. Let me give you two examples. So suppose you take f of x equal to x squared. This here is a polynomial. We know how to compute the derivative of any polynomial.

This is just using the Leibniz rule-- I mean, using these rules. In this case, you would only need the Leibniz rule together with-- we knew what the derivative of the function f of x equal to x is. So in this case, the derivative of this here is 2x. And so this means that x is equal to 0 is a critical point. It's actually the only critical point.

And so you see, in this case, this is how the function looked like. This is how the function looks like. Here, you have 0. Another example is that if you take another polynomial, where instead of x to the power 2, it's x to the power 3, then the derivative of this function is 3x squared.

And so you see again that x equal to 0 is a critical point. The derivative here, if you plug in 0, you get 0. So this is a critical point. But in this case, the graph of the function for this example here, the graph, it looks like something like this. It becomes very flat here, like this.

So 0 is a critical point, but it's not like it has a local maximum or minimum. OK. So the next thing I want to do is, I want to now apply it. I want to apply this simple observation. And I want to apply it to get a consequence, a simple consequence. And then I want to apply that simple consequence to get a lot of other consequences.

So the first thing I want to do is what's called Rolle's lemma or theorem. So this is the following, that I have a function f from a to b into. R.

And I'm assuming that f is differentiable. And I'm also assuming that the value initially and the value at the end, that those two values are the same. Sometimes one set those value equal to 0, but they just need to be the same.

And then the claim is that then there exist. a x0, where is strictly between a and b. And the derivative at this x0 is equal to 0.

And so the idea is just very simple, that you have a function. And the function initially-- this is set a. Here, you have b. The function initially and at the end end up at the same value.

So this here is f of a, and that's the same as f of b. And then who knows what happened in between? Maybe like this. And the claim is that then there's always some point where the derivative is 0. In this case, it looks like that those two would be the points where the derivative was 0.

And the proof is just this proof that there's three cases to consider. One is-- I mean, you don't really have to divide it up by this, but let me divide it up into these three cases. One is that the function here is constant, everywhere equal to the same.

Well, if the function's constant, then the derivative is 0. We have already proven that. Case 2, somewhere, there exist x such that the value at this x is strictly bigger than the value at the endpoints.

And then case 3-- there exists an x such that the value at this x is strictly less than the values at the endpoint. It's clear that one of these three cases happen, right? Of course, it could be that case 2 and 3 could both happen. But at least one of the three cases would happen.

In case 1, we already know that the derivative is 0 everywhere. So let's just look at case 2. Case 3 is similar. So if there exists some x where the value-- so in this picture here, this here is maybe x. The value here is larger than at the two endpoints.

Well, by the extreme value theorem that we proved earlier-- just because if a function is differentiable, it's also continuous. By the extreme value theorem, there exists somewhere where, actually, the maximum is achieved. And since somewhere in between-- in this case here, the maximum is not achieved at the two endpoints. So the maximum must be achieved at an interior point.

But now, it's not just a local maximum. It actually is a maximum. But already, if it was a local maximum, the derivative would be 0 at that point. So this proved case 2. And case 3 is similar. So this is Rolle's theorem Now, Rolle's theorem have then a bunch of applications. And the first one is to what's called the mean, value theorem.

And there are two versions of the mean value theorem. One is the actual mean value theorem. And the other is like a fancy version of it, where you have two functions in play at the same time. And that's called the Cauchy mean value theorem. So first consequence of Rolle's theorem.

And this is the mean value theorem. OK. And so the first one is the actual mean value theorem.

And the actual mean value theorem says that if you take a function f from a to b on some interval like this and the function is differentiable, then there exists an x, where x is strictly between a and b and so that the derivative at x is equal to the difference quotient, where you do the difference quotient at the two endpoints.

So this here is this side of it. That's what you often call the mean. And the mean is equal to the derivative at some intermediate point. This is what they say. And so this seems a lot better than Rolle's theorem. But it's an easy consequence of Rolle's theorem.

Because if you look at-- so suppose you take the function-- so proof. Suppose you take the function h of x equal to f of x minus f of b minus f of a divided by b minus a times x minus a.

Suppose you're looking at this function here, h. Then observe first that-- so this one, this function is differentiable. This is just a constant. This is just a linear function. a is a constant also. So this is differentiable. This is a constant. This is differentiable. And we're subtracting it from a differentiable function. So h is differentiable.

h of a is f of a. And then this thing here, minus f of b minus f of a, b minus a-- and then this here is a minus a. So this is 0. So this is f of a. h of b is f of b minus and then f of b minus f of a, b minus a.

And here, you plug in b, so b minus a. And so you see that these here cancel out. So you are subtracting this. But this is a minus minus, so you're adding. So this is f of a. So you see that for this function, the value at a and b are the same. So now we apply Rolle's theorem.

So now we have, by Rolle's theorem applied to h, then we have that h is equal to 0-- the derivative of h is equal to 0 for some x, where x is strictly between a and b.

This is what Rolle's theorem. This is because this function, h, has the same value at a as at b.

Now, what is the derivative of h? So the derivative of h is the derivative of f. And, again, this is just a constant. So we have to keep that minus f of b minus f of a, b minus a. So this is just a constant. And then derivative of this is just 1. So this is what the derivative is.

And so you see that-- so we have this x, where this derivative is 0. So this here is 0. But if this here is 0, it means that this thing here is equal to that. And so that was the mean value theorem. So this proved the mean value theorem. OK. There's a second mean value theorem. And the second one is the Cauchy mean value theorem.

And so this is a little more fancy. So the Cauchy mean value theorem-- so this is the following, that you have two functions, f and g.

So we have f and g on some interval from a to b into R. And they are both differentiable. They're both differentiable.

And then the claim is that there exists a x-- let's call it x0. Maybe I should have called it x0 up there, it's just one. There exists x0 with x0 strictly between a and b.

OK. And so it's something about that the derivative of both functions kind of behave like the mean value theorem at the same time.

So it is that the derivative-- I will explain this in a second-- that the derivative at this point times g of b minus g of a is equal to the derivative of g at x0 f of b minus f of a.

So this is the Cauchy mean value theorem. So let's say that for this here. And let's just look at that expression here before we prove the Cauchy mean value theorem.

So as a consequence-- so observe that, if g of b is not equal to g of a and g prime at x0 is not equal to 0-- if you have that, then you could divide-- you could have this expression.

Let's say that-- then, by just dividing this here on both sides and this here on both sides, you would get that f prime of x0 divided by g prime at x0 is equal to f of b minus f of a g of b minus g of a. You would have that.

Now, I could think about this thing here as f of b minus f of a divided by b minus a. So I could think about it like this and then times b minus a, g of b minus g of a.

And so you see that this here, it's something about-- the actual mean value theorem say that this thing here, that there would be some point, x0, in between that was equal to this. That would be the mean value theorem applied to f. And the mean value theorem applied to g would be that there exists some x0 so that 1 over this is equal to this thing here.

And the Cauchy mean value theorem say that you can arrange so that both of these things happen at the same time. So this seems actually pretty good. You can do it at the same time. It's a little bit surprising, in a way. But the proof is quite trivial. But it's useful.

So the proof is the following, that you just look-- so this is what we want to prove. So you just define-- and the exterior is unknown.

So if I take this expression that I want to prove, if I move this thing here over on the other side, then I would get-- so in this case, I would define h of x to be f of x g b minus g a minus g of x f b minus f a.

So I would define it like this. And you see that when you now take the derivative of h, you just get this here. This is because everything else is constant.

So this is what you get. And you want to find a x0 where you want to find-- so want to find x0 where this thing here is 0. To find such a x0, it's the same as saying that you find an x0 over here.

So now let's see. What is the value of h at a? If you plug in a here, then you get f of a, g of b minus g of a minus g of a f of b minus f of a.

And so now you see here, I have this one here. I have f of a times g of a. That comes with a minus sign. I have the same term over here, but that comes with a plus. So those two cancel out. So I just get f of a g of b minus g of a f of b. So that's what I get there.

And now I look at h of b. So if I plug in b up here for x, then I get f of b g of b minus g of a minus g of b f of b minus f of a.

And now you see this when you multiply this by this, that term, f of g of b, you have this over here also, but it comes with a minus sign. So those cancel out. And so you're getting a term-- the only term that lives is this one times that, so minus g of a f of b. And then you get this term, minus minus, so plus g of b, f of a.

Right. That's right. Sorry. And now you see that this here is the same as that. And these two here are the same. So we from this, we conclude that h of a is equal to h of b.

So again, it's satisfying the requirements of Rolle's theorem. And so there is this intermediate point where the derivative is 0. But if the derivative of h is 0, that's exactly the point we wanted to find, where you have this. OK. And so, again, the Cauchy mean value theorem is sort of a stronger statement. It's a stronger statement because it allows you to have two functions.

And then you know this point, where, to some extent, the mean value for both functions are satisfied. That's the idea. But it's an example of where, actually, the proof, it, in fact, worked more generally than what you used it for initially. In the initial, you just used it for a single function to prove the mean value theorem.

OK. So now we want to prove-- so now we have a number of consequences of this.

And the first one-- let me start with another consequence of Rolle's theorem. And then I come back to the Cauchy mean value theorem.

So Taylor, Taylor expansion-- so the idea here is that you have a function, f. And you have a function. Let's say that it's a function on R-- doesn't need to be defined on all of R-- into R.

And f, I'm assuming-- so let me do a special case of this, so special case first. So I assume I have a function like this and that it's differentiable.

And then I'm assuming that the derivative-- so the derivative is now a function from R into R. And I'm assuming that the derivative is also differentiable. And the idea is now that I'm looking at the function. So the idea is like this.

Again, in a minute, I'll prove a more general statement. But I just proved this one first to make the idea clear. So the idea is that you want to say that f locally can be approximated by a polynomial.

So that's the idea of Taylor expansion, that if you take a function-- oops. If you take a function like this, some function, and you take a point, then in a neighborhood-- let's say that this is at some point here, x0. Then in the neighborhood of this point, instead of looking at the function, you get could get a very good idea about how the function behaves in this small neighborhood if you look at a certain polynomial.

And the higher degree the polynomial, the better the approximation is to the function. And so we'll just do a simple version of this first, where you're looking at the-- so in this case, I would look at the-- so we have f here.

And I'm interested in f. So we will be interested in f near x0. And then you're looking at what's called the Taylor polynomial at x0.

And the Taylor polynomial, it depends on how many times you're willing to take the derivative. And so in this case, I just look at the first Taylor polynomial. And now it is a Taylor polynomial at x0. So this thing here is, by definition, the value of f at x0 plus-- and then the derivative of f at x0 minus x0.

So this here is the first-order Taylor expansion of-- so this here is often called the first-order Taylor expansion at x0.

Now, what you're interested in is, you're interested in showing that this polynomial here-- so, again, the idea is that f of x should be almost equal to P1 of x if x is almost equal to x0.

So near x0, the function is very well approximated by, in this case, this linear function. So now the key here is, how do you estimate how well it approximates?

So the statement is then that-- so you have here x0. You have here x. And the claim is now that there exist some c between x0 and x, so strictly between-- so between, which means that it's strictly between.

Let me just write that. x could also lie on the other side. It's so obvious what I mean, strictly between also. So there exists some c in between these two so that f of x is equal to this Taylor polynomial plus and then the second derivative.

So this is the derivative of the derivative. So f2 here is the derivative, and then you take the derivative like that. And that is evaluated at c and then divided by 2. And then it's x x0 squared. And then in a minute, I will just do the general case of it.

OK. And so now, how do you prove this? So what you do is-- so let's try to prove this.

So I define a M. We define a M. M is just some real number. And it's defined so that f of x-- I wonder if I should-- in my notes here, I call it, just because I thought it was clearer-- and maybe I will do this here also.

I'll call this point here a and this here b. And then c here is somewhere in between. So this means that this thing here is b, and this is b. And this here is then b. And this here is a. It's just to be clear that somehow, you're kind of fixing x.

So to prove it-- so we look at this thing here. And you look at thing here. And we subtract P1 of b. And then we define M so that you have that this thing here is b minus a squared. So define M by this thing here.

This here just gives you some number. a and b are different. So this here is nonzero. M is just given by-- that is this difference here divided by this here times 2. So that's how we defined M. But now we need to show-- if you do it like this, it's like M is OK.

It's just, by definition, you have-- so by definition, we have that f of b again is equal to P1 of b plus M 2 b minus a squared.

But we need to prove-- we need to show that M here is the second derivative of the original function at some intermediate point. That's what we need to prove.

And so you're looking at-- so now we're looking at the function h of x equal to f of x minus P1 of x minus M/2 x minus a squared.

So if we're looking at this function-- well, if we're looking at h of-- let me do b first. So if I do b, then this here is f of b minus P1 of b minus M/2 b minus a squared. But M was defined so that this thing here is equal to that.

So by definition-- this is by definition of M-- I have that this thing here is 0. If I'm now looking at h of a-- well, this thing here-- here, you get f of a here, you get minus P1 of a. If you plug in a, here then this is 0. So this is what you get.

But remember how P1 was given. P1 was here. And we now do it at a. So this here is really a. So you see that when you plug in x here, this thing here is 0. So you just get f of a here. So this thing here is f of a minus f of a. So that's 0.

So you see that if you're looking at this function here, then at a, it's 0. At b, it's 0. So by Rolle's theorem, there exist a c1 between a and b such that h prime of c is equal to 0.

Sorry, this is the c1. This is not the c we're looking for. The C we're looking for is something where the second derivative is something. But now you see that now you have this function, h. And here, you have a. Here, you have c1. And then b is further away, like this.

And you have that in this case, the derivative is 0 at a. And the derivative-- sorry, not at a. But the derivative at this point here, h prime, is equal to 0. Let's calculate what the derivative of h is at a.

I want to apply Rolle's theorem a second time. But in order to apply Rolle's theorem a second time, I really need the derivative of h at a to be 0. But let's see what it's equal to.

And so, again, we have h up there at the top. And so this means that h prime is equal to f prime minus-- and then the Taylor polynomial here was the Taylor polynomial. When you take the derivative, this is just a constant. You're getting f prime a. This is all it gives when you take the derivative of this here.

And then you're getting minus. And you're getting M x minus a. That's when you take the derivative of the last term. So this is what you get. And this here is at x. But so now h prime at a, this is f prime at a minus f prime at a. This thing here is actually 0. So you see that this whole thing is 0.

So what we have is that we have that h prime at c1 is 0. And this is also equal to the derivative of a. So we can apply Rolle's theorem a second time.

So by applying Rolle's theorem a second time, we have a c2. And now this c2, that's really going to be our c in this case.

So we have a second c, where the derivative of the derivative. S-- The second derivative of h-- at that point is equal to 0. And this one here, remember, this here is-- when you apply Rolle's theorem, you started with a. You had here b. Then you got c1 that was between these two. That was c1. But now you're looking at h prime in this interval. And so the second one is lying here.

So it's certainly between a and b, but it's actually even between a and c1. Now, what does this here mean? So we hope that this is the point we're looking for. And so to calculate the second derivative-- so this here was the derivative of h.

So we just need to take another derivative of h. So if we do this, then we get that-- so h double prime of x at some x is equal to f double prime at x. And then you see this thing here. This is just a constant. This is a linear function. So the derivative of this one here is 1. So you're getting this minus M.

And now we know that h double prime at this intermediate point, c2, which we also call c, that's equal to 0. So this means that f double prime at the c minus M is equal to 0. So this means that M really was the second derivative of f at this point.

So this proof's sort of clever. The proof seems very clever. But if you think about it-- so it's applying Rolle's theorem twice. But it's, in a way, the only thing it could be.

But anyway, OK. So that was Taylor. This was just a special case of Taylor expansion. Let me state the general case. So for proving the special case, we applied Rolle's theorem twice. But you can actually apply Rolle's theorem k times.

And so if you apply k times, then you're getting the k Taylor expansion. So this is the general case of the Taylor expansion, so general case of Taylor expansion.

So the general case is that you take a function on R into R. You assume that f is differentiable but also that f prime is differentiable and all the way up to, say, the k-- right.

You assume that it's differentiable, and all the way up to the k-minus-1 derivative is also differentiable. So this here exists. So all of these exist. All these derivative exist.

And then the claim is that-- so now the Taylor polynomial, the k Taylor polynomial-- or sometimes called the Taylor polynomial of degree k minus 1 is that you're looking at-- oh, sorry.

This here should be the k minus 1 Taylor polynomial. So this is where you write-- so you take-- again, you have here a, and you have b. Let me write it with a and b instead of with x, because I think it's a little clearer. So the thing is that the polynomial is where you say you expand around a.

And so the polynomial, this is the k-minus-1 polynomial. For that, we write it in terms of x. And this is the polynomial where the constant term is the value of f at a. The coefficient to the first-order term is f prime at a and then x minus a.

The coefficient to the second is the second derivative at a. And now it's divided by 2 x minus a squared. And then this next one is f triple prime. And now it's really 3 times 2 times 1 times x minus a to the power 3.

And then you continue like this up to the k minus 1 derivative and then this k minus 1 factorial, x minus a, k minus 1. So this is the k minus 1 Taylor polynomial.

And so you observe that you can just write it-- so it's typically written like this, that if the i derivative in a divided by i factorial x minus a to the power i-- and i is now summing from 0 to k minus 1. This is by definition equal to that.

So that's the Taylor polynomial, the k minus 1 Taylor polynomial. And now the general case of the Taylor expansion is then-- so the general case of the Taylor expansion is then.

So this is like the theorem, Taylor expansion. So you have again this function on R into R. And you have that all of the derivatives up to and including the k derivative, they all exist.

And then you have that f of b is equal to the k minus 1 Taylor polynomial at b plus-- and then this term is usually called the remainder term.

And the key point here is that it's easy to estimate the remainder term because there's an expression for it. This is the k derivative of f in this point that is an intermediate point divided by k factorial b minus a to the power k and where c here is lying between a and b.

And again, the idea is that you have a function, any function. It needs to be differentiable k times. But then, actually, it's very close to a polynomial of degree k minus 1. And you can estimate it, how close it is, in terms of this term here.

And so the proof of this-- so the proof is really just what we did, but let's just go over it again. We only did the first, where we did it in the case where k was equal to 2. And now let me try to do this in general.

So you're looking at f of b minus-- we have this Taylor polynomial, k minus 1 Taylor polynomial. And you define M by the equation, that this here is m over k factorial times b minus a to the power k.

So this here is how M is defined. b is not equal to M. So this is just some positive. It's actually a positive constant. And so you can just write-- so M here is given as the difference of those two. And then divide it by this and multiply by k factorial. So this defined M.

And now what you do is, you then define h as before or similarly to before. h is then going to be f of x minus P k minus 1 of x minus M over k factorial x minus a to the power k. That's how you define h.

And by definition of M, by definition of M, you have that h of b is equal to 0. This is another way of saying how we define M.

So now h of a-- well, h of a is just f of a minus p k minus 1 of a minus M over k factorial x minus a to the power k. So this one here is clearly 0. a minus a 0. So that's clearly 0.

And this thing here, that's this polynomial. I wrote it up here. So if you plug in here a, you see this term here and all the other terms drops out. It's only this one that lives. So this thing here, this is the same as f of a. And so you have that the whole thing is actually 0.

So because we chose M this way, we see that now Rolle's theorem can get started. And Rolle's theorem can get started. And this means that there exists a c1 between a and b where the derivative of h is 0.

Now you check that-- if you're looking at this function again and take the derivative at a, the derivative at a is also 0. In fact, the whole idea is that all these derivatives up to the k minus 1 derivative of-- if you're looking at the k minus 1 all the way up to the k minus 1 derivative of h, they are all 0 at a.

So you can repeat the Rolle's theorem So first here, this thing here together with this allows you to find this c1 strictly between a and b, where the derivative is 0. Then you observe that the derivative of h at a is 0.

This then allows you to find a point between a and a second point, c1, which was less than b, where the second derivative is 0. Now you use that this second derivative of h at a is also 0.

So you apply Rolle's theorem one more time. Each time you apply Rolle's theorem, the c's is moving to the left towards a. And so you apply it. You apply Rolle's theorem k times. And once you do this, you get eventually that this m here is the derivative of the original function at this point.

So I wrote this out. Now, I did the special case, and I try to explain this here. I wrote it out in the notes. It's written out in the general case only. But on the other hand, it's written out completely there. Any questions?

There's two things we didn't get to. There's these two rules about how to evaluate fractions that uses the Cauchy mean value theorem. And I'll start with those. Those are in the notes, but I will start with that next time. Yeah?

STUDENT: So M is like [INAUDIBLE] derivative?

TOBIAS COLDING: You do up to k derivative. It's when you do the k derivative, then that's when you get the c that you're actually looking for. And so in the special case I looked at, k was equal to 2. And the Taylor expansion was a linear-- so it was a polynomial of degree 1. So k minus 1 was 1. And so in the special case, we applied Rolle's theorem twice. And in this case, we do it k time.