[SQUEAKING]

[RUSTLING]

[CLICKING]

TOBIAS COLDING: Last time, one of the things we did were we proved true mean value inequalities. So we proved true mean value inequality. And let me recall those.

And the first one was that-- so if you had a function f on some interval from a to b into R and this function f is differentiable then it said that there exists some x0 with-- x0 between strictly between a and b so that the derivative at x0 is equal to the average of the function-- or the average of the function at the endpoints. So this is the average on the endpoint divided-- so it's average over the length of the interval. So this is what the mean value theorem said.

And then there was a second mean value theorem, and that was the Cauchy mean value theorem. So the Cauchy mean value theorem. And that was where you had two functions, f and g, on this interval into R. And they are both assumed to be differentiable.

And then again, it was that there exists some x0 strictly between a and b. And if you take the derivative at x0 of f times the value-- the difference in value of g at the endpoints, then this is equal to the derivative of g at x0 times the difference in value of f at the endpoints.

All right, so that was the Cauchy mean value theorem. So it was a little bit more complicated than the actual mean value theorem because you had two functions at the same time, and the point here was that you can arrange this exterior to work for both points-- for both functions.

Let me just make an observation, that if-- so if we're looking at-- so the Cauchy mean value theorem seems like a more elaborate version of the mean value theorem, and it really is because if you suppose that g here was a function that is equal to x-- so g of x is equal to x, I have f here on the integral from a to b-- this is, again, differentiable.

Well, if the other function was just this rather trivial function g, of course, g is also differentiable, and we even know that the derivative of g at all points is 1. So if you are now applying the Cauchy mean value theorem, then you say that by the Cauchy mean value theorem, there exists some x0 with x0 strictly between a and b.

And so that the derivative of f-- so that's whatever that is. That's the derivative of f at x0. The value of g at these two endpoints where g is just a function of x-- so this is just b minus a. And the derivative of g is 1. And then it's the difference in value of f at the endpoints.

And so now, if you just divide by b minus a-- so this, of course, implies that f prime of x0 is equal to f of b minus f of a divided by b minus a. So in other words-- and you see, this here gives you-- this gives you the more standard mean value inequality. So in other words, the Cauchy mean value theorem is more general than this mean value inequality because you recover this-- if you just let one of the two functions, it'd be just the trivial function x. OK, so that was just an observation.

The second thing I want to talk about is something we didn't get to last time. And these are these-- L'Hopital's rule. And I'm just going to state one version of it. There are two versions, and they're both in the notes, but I will just state one of them.

And so this is-- so you have-- so the situation here is that you have two functions, f and g. And they are defined on some interval, say from a to b into R. The interval does not include the endpoint. So these are differentiable.

And then we assume that the limit here-- we assume that g of x is-- and not only g of x, but also g prime of x, are not equal to 0 for all x. And then we assume that the limit here of g of x as x goes to a, that this here is 0, and the same holds for the function f. So this here.

And then what you're interested in is-- what you're interested in is that you're interested in looking at f of x divided by g of x. So this here makes sense, g doesn't vanish. So this here makes sense. And you are interested in looking at this fraction here as x goes to a. If x goes to a, then the denominator goes to 0, and the denominator also goes to 0. And the question is, does it have a limit or not?

And so there's one more assumption here that you need, and that is that if you're looking at the-- so assume also that the limit here as x goes to a of f prime of x divided by g prime, assume that this exists and is equal to some L.

And then the claim is that then-- so the claim here is that this fraction here, f of x divided by g of x, that this does have a limit as x goes to a, and that limit is actually L. So this is one version of this rule. And there's another one, very similar. OK.

And so we want to prove this. So we are looking at-- so we want to apply the Cauchy mean value theorem to prove this theorem. And so we're looking at an interval here. So you're looking at, say, y. So here, you have a; here, you have x; and then here, you have y; and here, you have b.

And I'm now going to apply-- so we will apply the-- so apply the Cauchy mean value theorem. But we are applying it on this interval here. Between-- so y here is less than x, but is strictly bigger than a. And this is x is strictly less than x.

So, the function was differentiable on this open interval-- so of course, differentiable on both functions, differentiable on this closed interval. So we can apply it-- and so we get that f prime at some intermediate point x0, times g of x minus g of y, this is equal to g prime-- again, x0 here is some intermediate point, it lie between these two. g prime at x0 times f of x minus g of x-- g of y, sorry, like that. This is just applying the Cauchy mean value theorem.

Sorry? What did you say? Oh, yeah, yeah, yeah, sorry about that. Thanks. Yeah, thank you. Right. And so now, I just divide-- of course, this here is supposed to be a non-zero, so I can just divide. So we have that f prime of x0 divided by g prime of x0.

And then I will-- I want to divide by this, but I just need to make sure that this here is-- that I can divide by it, that is not 0, this thing here. So I have this.

Now, think about x as being fixed and do this for all y. If I do this-- if I do this for y where y is very, very close to a, this here becomes really, really close to 0. This here is fixed and it's not equal to 0. So this here is very close to 0, this is not equal to 0. So this here, if y is really close, if x is fixed, y is really close to a, this is definitely not zero, so I can divide by it.

So, for x fixed and y sufficiently close to 0, then you have, then, g of x minus g of y is not equal to 0. And so I get that f prime at this intermediate point divided by g prime at the same intermediate point is equal to f of x minus f of y divided by g of x minus g of y.

And now you see that's y here goes to 0-- so x is fixed still as y goes to-- sorry, not to 0, but to a, this here goes to 0 by assumption, this here goes to 0, so the whole thing goes to f of x divided by g of x.

Now the thing is that if you-- so this is for all x fixed. All right, so now if you-- so now, we want to use-- so now, since f prime at, say, z divided by g prime at z, since this fraction here, as z goes to a, this is assumed to go to L.

So this means that given epsilon greater than 0, there exists a delta greater than 0 so that if z here is less than a plus delta, then this fraction here-- then f prime of z minus g prime of z is-- minus L is less than epsilon.

So now coming back here, if you were starting with-- so this here, again, was for all fixed x. You had this statement. And again, x0 was between y and x. So now if you-- so if this x here-- if x is-- is delta close to the left endpoint of the interval, then you would have that x0.

But x0-- remember that x0 is lying here. So since x0 would lie here, this means that also, x0 here is in this interval here. And so this means that this thing-- this number here you get is L up to some-- it's L up to some error of epsilon. So this here is almost equal to L. And so this means that this limit here is almost equal to L within an error of epsilon. So this proof-- this version of L'Hopital's rule.

And that's the second version, and that's in the notes, but I'll skip it, actually, for now just because there was-- from the last lecture, and it's very similar to this. Yeah?

STUDENT: Why did you deduce the g x minus g y is not equal to 0?

TOBIAS COLDING: Sorry? Why--

STUDENT: Why would you deduce x that g x minus g y is not equal to 0?

TOBIAS COLDING: Yeah, OK. So this is because g of x is supposed to be not equal to x for all x. And now I'm fixing x. So this here is not 0. Then I'm letting y go to a, but then this here goes to 0 by assumption. So you see that if you're fixing x and you just take out y sufficiently close to a, then this denominator is definitely not 0. You really want to let this-- you want to let y goes to a because then this here is eliminated. This is not part of the statement, that g of y.

And so, again, as long as we started with an x that was sufficiently close to y-- if x was in this interval, then x0 is even further to the left, so this is also in this interval. We don't know exactly what x0 is, but it is in that interval, and so this means that this is strictly smaller than epsilon.

OK. So another thing we talked about last time was this with the Taylor expansion. In that all of these things really came out of this broader theorem. And that was also the case for the Taylor expansion.

So you have a function f-- so this is-- just to remind you about the Taylor expansion. So you have a function f. And let's say that it's defined on some interval here into R. And the function is-- so f here is k times differentiable. So this means that f prime exists, f double prime up to the k derivative. So they all exist.

And then the statement was-- so then you looked at the k minus 1. This is the k minus 1 Taylor polynomial. And it was given by that k minus 1 of x is equal to-- and now you're summing here from i equal to 0 to k minus 1. And you take the i derivative-- this is-- sorry, this is the i derivative of f, and you do this in a, and you divide by a factorial, and then you're looking at x minus a to the power i. So this was the k minus 1 Taylor polynomial.

And then the Taylor expansion just told you that f here at some x was equal to P k minus 1, this Taylor polynomial plus. And then this is what is usually referred to as the remainder term.

And this looked very similar to one of these terms. It was the k derivative of f, but then it was evaluated at a point c, divided by k factorial times x minus a to the power k And c here was supposed to lie between a and x. So that was-- this was the Taylor expansion theorem.

And so the thing is that-- so why is this so useful? And it is extremely useful in many different settings, also in much more general settings than what I'm describing here.

And that is because that if you take a function here, then-- and it may be a very complicated function, but this Taylor expansion theorem tells you that it is close to-- if you want to evaluate it at some x, you can evaluate the Taylor polynomial-- that's, after all, just a polynomial, so that's easy to evaluate. And then you get within this remainder term of the right answer. So that's how it always is used.

And so the question is now-- the natural question is, how well does the Taylor polynomial approximate the value of the given function? Right, so how well does the Taylor polynomial approximate the function?

And of course, to understand that, you need to estimate the remainder term. Let me write the remainder term here. So the remainder term-- right, so the remainder term is-- so I write this here as-- so it really depends on x also, but I will just write it like this in here.

So by this, I mean the k derivative of f in c divided by k factorial, and then x minus a to the power k. So again, it depends on really also-- of course, it depends on a, but it also depends on x, and it's this thing here. This thing-- the catch is, that you don't really know what c is. The only thing you know about c is that c is lying between a and x. So this is essentially all you need-- all you know about c.

So now, let's look at some examples. Let's look at two examples. So the first one is like a promising example, and the second one is much less promising. And so the thing is that how well it approximates really depends on the function. But the first one is quite typical, though, example.

So suppose that the function f of x is the exponential function. And then we have that the derivative is just the exponential function itself. So this means that if you take the derivative k times, then you just get the function itself. So this means that the k minus 1-- which means that the k minus 1 Taylor polynomial, this is just the value of the function-- I'll just write it out here. So this is the value, and now I do it-- I do it for a equal to 0.

And I'm assuming that b here-- so I'm thinking about-- I will think about this function here f of x on this function, the exponential function. I'm going to think about it on the interval, say, from minus 1 to 1 into R. Maybe I'll just take it from 0 to 1 into R.

OK. And the Taylor polynomial-- so I'm going to-- so the Taylor polynomial, one, also talk about that this-- the Taylor polynomial is expanded around a. So we talk about the Taylor polynomial. So we talk about Taylor expansion around a. And this just refers to that when you plug in a here, well then, of course, you have exact equality. And that all of the derivatives are evaluated at a.

So it's really-- like the Taylor polynomial, it's only a good approximation near a. So I want now to Taylor expand near the theorem. And so the value here of the function, this 1 at 0, the derivative is-- so the derivative is also 1 because it's just a function itself.

And so this is-- so you get x, and then you get-- the next element in the Taylor expansion is x squared over 2 plus x squared over 3 factorial. And then x to the fourth over 4 factorial, et cetera, up to x to the power k minus 1 over k minus 1 factorial. So this is the k Taylor expansion, or k minus 1 Taylor polynomial, around 0

And so now, if you are interested in evaluating P, you want to say, how well does this here approximate the original function, the exponential function, so you want to bound this thing here. So you want to have a bound for that. Well, the thing, again, is that f of x, you know that f of x is equal to this Taylor polynomial, and then plus this remainder term in the c.

Now c, again, has to lie-- c will see will lie between a and x, and in this case, a is 0. So it means that c here-- so this means that c here is between 0 and x. And the remainder term here, this here was equal to-- that you took the k derivative of the function. So in this case, the k-th derivative of the function is the function itself. So this here is the value of the s derivative, and this is at c. This is the k-th derivative of the original function. And then you have to divide by k factorial, and then it's x to the power k like that.

And so you see that if x here is now between 0 and 1, this means that this thing here, this remainder term here, the absolute value is-- so c here is less than 1, so this means that this thing here is less than e. So it's less than or equal to e divided by k factorial. This here is bounded by-- since x is bounded by 1, this is bounded by 1, so you're getting this inequality here.

And so you see that if k is very large, this is a tiny number. So it's a really good approximation. I mean, you don't even need to make k very large because it's exponential. Not exponential, it's factorial. Yeah. Right. OK. So that's a kind of encouraging example. Let's say that the Taylor expansion is a very-- in this case, it's a very good approximation to the function.

Another example that is less encouraging-- and so again, the moral is that how well the Taylor expansion approximate the function depends on the function. So suppose that the function here was a function that was 0 here when x is less than 0-- less than or equal to 0. And when x is bigger than 0, then it is e to the minus 1 over x squared.

So you see that when x becomes very, very small, you are looking at 1 over the square. So this here is huge. But now it's with a minus sign, so this means that this thing here is incredibly small. And so it's clear that this function here is continuous at 0. It's clearly continuous everywhere else. But actually, using in this argument, it's easy to see that this function f here is infinite differentiable. So you can differentiate it infinitely many times, and even f0.

And by its end, of course, the derivative is-- so it's an example of a function that looks like this. It's that 0 until 0, and then it goes up like this. And so the derivative at 0 is 0. Any derivative, however many times you take the derivative of this function-- so this thing here, if you take any derivative and you evaluate it at 0, you always get 0.

So this means that if you're looking at the Taylor expansion at 0-- so you're looking at this k minus 1 Taylor polynomial, this polynomial is-- all of the coefficients in the polynomial are 0. So this polynomial is just 0. So this means that if you're looking at f of x where f of x is actually equal to this remainder term, so it's somehow-- the Taylor polynomial is too trivial. It doesn't really help you all that much. So, it's equal to this remainder term. OK. Yeah?

STUDENT: Is there a way of-- so the remainder-- like, boundedness helps us figure out how well the expansion approximates the actual function. Is there a way to determine how far from the point you're approximating? Like a or 0 in this case, you can go to a certain endpoint?

TOBIAS COLDING: Yeah, right, that's right. I mean, so you are asking-- so you're asking-- suppose-- and you definitely could do this. You're asking-- you're thinking that you have to have a-- you see a slightly different situation than here, but closely related. That you have some a here. And you have a function f.

And you want to-- and you are willing to maybe look at-- you're willing to take some fixed large k. And then you're asking-- what I'm interested in, how well does this-- how well does this function, this polynomial approximate this? But I want it within an error, certain error.

So there's a given epsilon, and you are interested in how far away can you go from a so that this here is less than epsilon? And so this is-- so this becomes a question of this remainder term. So this becomes a question of this remainder term. When can you always say that this remainder term is less than epsilon?

And so in this particular case, you would have to do it on a case-by-case basis. But in this particular case, the remainder term is-- If this interval here, in that example over there, if you have here 0, this was a in our example, and then you want to go out here to some x, and now you know that C here is between 0 and x, but you don't really know anything else, so then it's like this remainder term, you will have this remainder term in this example adds this c here. This, again, was equal to this thing here. Over k factorial.

And so you would have-- that you would want this thing here, Rk c, You. Would want this to be less than epsilon. And so you would have that this thing here-- the only thing you can really do is that you can say that c is less than or equal to x.

STUDENT: And we solved for c--

TOBIAS COLDING: Yeah. That whenever-- so you would do-- so what you would do here is you would say Rk c, this here has to be bounded by-- because x c is less than x. So this is e to the x-- you have to eliminate c. So you would do like this, k factorial times x to the power k. You have to eliminate c because you don't know anything really about c.

And so this here now has to be less than epsilon. So now you only have-- you have fixed your k, and so now you only have one variable and you can solve that. But you really have to eliminate c, yeah. OK. So again, so how well the Taylor expansion approximate depends on the function.

OK, the next topic is another important topic. And so this is integrals. And there are different ways of defining integrals. This here is-- this is certainly one of the-- maybe the most natural ones. And so this is what's called Riemannese. So it's after German mathematician Riemann, integrals.

And so the thing is that here now, you look at a function. You have a function on some interval from a to b into R. And this function f here is bounded. And the idea is-- so the idea is for integrals is-- so you have a function. And you want to be able to define-- so define an area below graph. So if this has a function, you want to be able to calculate what is this area.

Now, if the function was to-- if the function was to have crossed the x-axis, then this area here would count negatively and this area here would count positively. Anyway. So this is what you want to do. And you're always starting with a bounded function.

Now, to do that-- right, and so you want to define that. The only thing that you know for sure is-- so there's only one ingredient you have. So only thing you know at the outset. So the only thing is that if you take a right angle-- so if this length here is-- let's call it d, and this here is e, then you know that the area here-- so area-- is d times e. That's the only thing we have.

So now, how do we define the area below the graph? And to do that, we first need to talk about partitions. So partition. So suppose I have an integral here. a and b are real numbers. So I have integral here. Then I talk about a partition. And it's useful to have a notation for partition. So I call it P of this interval here.

This is finitely many dividing points. And the dividing points always have to start at a and end at b. So you have finitely many-- let's say k many. And you're looking at k plus 1 many. So the first one has to be a. And then you have x1 up to the last one, and that has to be b.

And so what you are thinking about is-- so you have this interval here from a to b, and you have these dividing points. And you are thinking about dividing the original interval into these subintervals. So these are the subintervals. So you divide the original interval into subintervals.

And the subintervals are-- I'll do it like this. I will do it where I take from-- it's always from one to the next. And I'll write it typically where I'm starting-- or I'm ending at i and beginning at i minus 1.

So I'm dividing it into these subintervals. So in this case here, in this case here where this is x0, there's going to be k subintervals. And you only always go from one to the next. So that's a partition.

And now we need a second notion related to this, and this is a subpartition. So subpartition. Subpartition. This is also sometimes called a refinement. So a subpartition is-- so if you have a partition-- P1 is a partition, and then P2 is a subpartition of refinement if P2 has all the dividing point of P1 and then some-- and then some. I mean-- right.

I mean, of course, it could be a trivial subpartition where P1 and P2 are the same. That's OK. But in general, it will have some more points-- also some more dividing points.

Now, right. So now, let me look at-- so what is, then-- so now, let's try to get closer to try to define this with the area below the graph. So again, we have a function f from a to b into R. We assume that this is-- so that is bounded. We have P here is a partition of this interval here. Then we defined a upper and a lower sum. Then we define upper and lower sums as follows.

So again, you have a given function that is bounded and you have a partition. And so we define upper and lower sums, and the upper and lower sums I will denote like this. It depends on the function, of course, and then on the partition. That's an upper sum. And the lower sum, similarly, I will denote like this.

And so how are these defined? So they are defined as follows. That you have-- here, you have your interval from a to b. Then you have-- this here is x0; here, I have x1; and here, I have maybe x2. And then this thing here is x3. So I take-- I look at one of these intervals. Well, it could be this one here. It's between-- in this case, it would be between x1 and x2.

And I would look at the function in this interval here, and I would-- so the contribution from this here would be that I would take-- so for the upper sum with respect to this partition, the one that comes from this subinterval here would be the sup of f on this subinterval here. So the sup of the function times the length of the interval like that. So this would be the value that-- in the upper that would come-- in a minute I'll write it out more clearly. But this would be what is coming from here.

And likewise, for the lower one, then the contribution from this little interval here would be the inf of f, and then times. And the idea is-- OK, so that would be the contribution. The idea, of course, is that you have-- the only thing we knew was this was the area of a right angle. And so you have this function here, and you have some little interval here between xi and xi minus-- sorry, xi minus 1 and xi.

And you're focusing on this little interval here. And you're looking at the function. The function was bounded. So there is a sup on this. And you try to look at the smallest right angle that kind of have the graph that fits-- where the graph fits inside that. And then you're looking at-- so this is the contribution from this little interval to p-- to the upper sum.

And then you have that you're looking at the largest right angle like this that fits under the graph, and that's in contribution to the lower one. And the idea is that if the function is reasonable, if you're looking at finite partition, then you should certainly get a better approximation to the area below the graph. I mean-- and if the function is reasonable, these two--

OK, so now what is, then, the-- let me just define this. So this thing here, it would be convenient to the length of this interval to denote it by this. So this is xi, xi minus 1. This is just the length of that interval.

OK. And so now written out, this thing-- and-- OK, let me make one other notation. Mi is the sup of f on xi minus 1 xi. That's capital Mi. And little mi is then the inf on the same interval like that. And so the upper sum for this partition, this is the sum from i equal to 1 to k of Mi times delta xi. That's the upper sum. And the lower sum is, then, the sum from i equal to 1 to k of Mi delta xi. So that's the upper and the lower sum.

The first thing you observe is-- the first thing you observe is a trivial fact. And it really should be the case because one is like what fits inside the graph-- below the graph, the other is where the graph fits below the right angle. So you clearly have just-- and this is just-- because this number here is the inf, this is the sup on the same interval. So you clearly have that L f, P is less than or equal to U f, P. So this is-- so this is clear.

OK. Let's look at-- let's look at a couple of examples. So just to make sure that this is totally clear. Let me look at a couple of examples. So the first example is-- so I take a function f of x equal to x squared plus 1. And my partition and the interval here, I'm going to think about this as a function on minus 2 to 2 into R. And my partition here is going to be minus 2 minus 1, 0, 1, and 2.

And, right. And so you have-- so let's see. So this here-- so if I draw the function, then this looks like-- if this is 1-- here, you have 1. Then it looks sort of like this. It doesn't-- more or less like this. And so this means that here and here, you have 2. And this here is minus 1. And here, you have minus 2.

So now, if you're looking at-- so, right. So this means here-- right, so this is my partition. So this means that x0 is equal to minus 2. x1, that's the next one, minus 1. x3 is the next one. That's 0. x4 is 1, and x5 is 2. And so, of course, all of these intervals here, they all have the same length in this case. So xi minus xi minus 1, this was what we call delta xi. This here is, for each of these, is equal to 1.

And now, we have that the-- and we have that the-- when you're looking between minus 2 and 2-- sorry, minus 2 and minus 1-- so in this interval here, the function is-- the function on this interval here is decreasing, so the largest value is achieved at this point here.

So M1, this is the same as the sup of f between x0 and x1. And so this is the same as the sup of minus 2 to negative 1 right of f. And the function was decreasing, so they lie on this interval. So the largest value is achieved here, and that's 5. So this would be 5.

And on the second interval here, so this is between negative 1 and 0. Still decreasing, so the value is-- the largest value is here where it's 2. So this here is 2. And M3, this is now between 0 and 1. So now it's increasing, so the largest value is over here. So that's, again, 2. And then the last one is, again, 5.

And then likewise, for the ones that are the infimum-- so the first one, M1, well, the function was decreasing, so the smallest value is over at that endpoint. There, the value is 2. So you're getting-- this here is 2. And then between negative 1 and 0, it's still decreasing, so there, the smallest is at this point where it's 1. So this is M2. And then M3, this is now-- here, the function is increasing, so it's still the value here. So it's still 1. And then the last one is between 1 and 2, and the infimum is the value here, so that's 2.

And so you see that in this case, the upper for this function-- for this partition, it's this-- all of the intervals have lengths. The small interval have length 1, so it's just the sum of these guys. So this is just equal to 5. If you will, it's 5 times 1, that was the length of the interval, plus 2 times 1 plus 2 times 1 plus 5 times 1. So in other words, you get here 14. And for the lower ones, it's, again-- this time I won't write times 1. It's the sum of those guys, and that's actually 6. So that's upper and lower sum.

Now-- and we had one observation that was an immediate observation, and this was this thing here. Now, what is useful is that you're starting with a partition, and then you have a subpartition. So the next one is-- so I take a function f on some integral from a to b, and this function here is bounded. And then I have P is a partition. P1, say, is a partition. And then P2 is a subpartition of P1.

And now, the thing that is useful is the following. And so the following claim, the following lemma here, that if you take a subpartition, then I claim that the value of the upper sum go down and the value of the lower go up. So this is the following.

So, lower of f times P1, this is less than or equal to the lower of f of P2. We already proved that this thing here, this one here-- this inequality we haven't proved yet, but this one here we have, that this is less than the upper one for the same partition. And the second claim is that this here is then less than the upper one for the first.

Now, why is this the case? This is also actually a trivial statement, but it's very useful statement. So imagine that you have your partition P1 here, and then you have the dividing points i minus 1, and here, you have i P2 is a subpartition. So there may be integrals like this where there are no further dividing points. But in that case, your the contribution for the lower and upper is the same for the two partitions.

But should there be one or more? Let's just say there's one. Then for the upper one, you would look at-- you would then get-- so you would get a new dividing point, let's call it y. And then you would get-- for the new one, you would look at the sup here on this interval here. But the sup on this interval here of the function is less than the sup on the larger interval.

So you would get a sup of f on the interval from i minus 1 to y of f. You get that. And then you would get times the length of the interval that's y minus xi minus 1. And then you would get a second sup, but this is now from y to xi of f. And then you would have multiplied by the length of this interval. So this here is then from xi to y.

So this is what you get from the new partition, the contribution from the second-- from the-- this is what? This is the contribution that you get from the upper sum for this partition P2 from this part. But now you see that this thing here, the sup over this little interval, well, that, of course, is less than the sup over the larger interval. So this is less than Mi, and I'll just keep this one here. And this one here is also a subinterval of this larger interval. So this is less than Mi again times xi minus x, y.

But now you can factor out the Mi. So you're getting Mi, and you're getting the sum of these two, y minus xi minus 1 plus xi minus y. And so you see that this is exactly Mi times xi minus xi minus 1.

And so you see that the contribution that we get here to the upper sum from this part here-- that was this and this, those two things, this is less than the contribution we get from the uppers sum, from the first partition. And so this proves this inequality, and this inequality is similar with just with the infs instead of the sups.

OK, so this is like a key point. Each of these steps is completely trivial, but together, it actually means that you can really define something non-trivial. OK. So now-- and this may seem trivial. Now we-- well, first of all, we define the upper and lower integrals. So we have upper and lower sum, and now we define the upper and lower integrals. So upper and lower integrals.

So the upper-lower integral-- so the upper here is denoted like this. We take a bar over this. And this is the upper-lower integral. So this is the integral, this is the function, and then you write this. And so this is simply defined to be the infimum over all partitions of the upper integral. Upper sum, sorry.

And so remember again that if you take a partition and you refined it, if you're looking at a subpartition of the given partition, then this number here goes down. I mean, it may stay the same, but it's certainly not bigger. So, this is-- so the idea here is that essentially, this number here you get by looking at finer and finer partition.

So, idea is that the infimum that this is a limit over finer and finer partition. So that's the idea, is the-- but the precise statement is that it's just the infimum. So this is the upper integral, and then likewise, we define the lower integrals.

STUDENT: Is there any assumptions that you have finitely many partitions? Or obviously--

TOBIAS COLDING: Right. But there is a key assumption is, that for each partitions, there are only finitely many divisions. But there's not a-- but there are infinitely many partitions. So it's really-- and so when I say that it's a limit over finer and finer partitions, you should really think about that in quotation marks. It's not very precise statement, but the precise statement is just that it's infimum over all partitions.

And so-- so-- I mean, you can make that precise, but then that requires a lot of stuff that we don't want to do here. And there's no reason to it because there's a perfectly reasonable way of defining it.

And so that was the upper integral-- so this is the upper integral. And then similarly, the lower integral, the lower integral-- and there, I'm just putting the bar below. And you define it like this. And this is now the sup because remember, that if you took a partition and then you take a subpartition, then this lower sum actually went up. So it's a sup of all partitions of this lower. OK.

Right. And so now we say that-- and then we say that-- so now, here's the definition. And again, this may initially seem we are wimping out, but in a minute, we will then determine when this is the case. So we define-- so f here-- so if you have a bounded function like this, then we say that f is Riemann integral.

So if the lower sum is equal to the upper sum-- so if this is the case. And in that case, and then the Riemann integral is this common-- so the Riemann integral is written like this. And this is just this common value. All right. OK.

Now, one thing I just want to remark is that we have-- you always have that the-- so one remark is that-- so note that-- so first of all, we had this trivial statement, that if you take a partition, then the lower sum is less or equal to the upper sum.

Then we had that if you took a subpartition, if this here was a subpartition of P1, then this lower sum here under a subpartition, the lower sum went up, but it's still staying below. So you have this inequality here. OK, now what is this useful for?

So as a consequence of this, we get that-- v or not the function is Riemann integral, you always have that the lower integral here of f dx is less than or equal to the upper integral. So this here is for all functions. And why is this the case? This is because if you take-- this here is the Soup of these lower partitions, but if you take any given partition, then this lower thing here is less than the upper one.

And so now, imagine that you take-- so if you take a lower partition-- suppose-- sorry. Suppose now that you take the sup over all partitions, then it means that there is one partition so that this value here is close to the sup. But for this one here, you have this inequality here.

And so in fact-- but in fact, if you take any other partition, you can think about-- maybe I'll explain this in more detail next time because I only have a minute. But it basically comes from-- this inequality here is coming from this inequality here. It will take just a few minutes to explain it clearly. So maybe I'll just leave it here.

But a key point here is that you have this inequality here. And this comes from that. But we return to that next time because to make it really clear, it will take a few minutes.

Now, the last thing I want to say is that-- so you may think that we wimped out of saying when this-- because we just define it, we just say that it's Riemann-integrable if those two things are equal. But then the question, of course, becomes, when are those two things equal? So that's the key question. So the key question is, when are they equal?

And we'll answer this next time where we will show that if f is continuous, then they're equal. And so anyways, I'll return to this. I will first-- we need, anyway, some more information about these upper and lower integrals, and I will first show why the lower one is less than or equal to-- the lower one is less than or equal to the upper one always. And then we will talk about some criteria for when they are equal. And then we will prove this for continuous functions.

STUDENT: Is that an if and only if or are there other--

TOBIAS COLDING: No. No, no, no. No. It's not. It's not-- it's-- yeah, it's not a total answer. It's a good initial answer, but there are many functions that are not continuous that are very worth-- that plays key roles, and they're still-- and we will also look at an example-- it's in the notes, but I didn't get to it, we look at an example of function that is not-- where you don't have this equality. And this is like an old.