The following content is provided under a Creative Commons License. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation or to view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at ocw.mit.edu.

Professor:

So, again welcome to 18.01. We're getting started today with what we're calling Unit One, a highly imaginative title. And it's differentiation. So, let me first tell you, briefly, what's in store in the next couple of weeks. The main topic today is what is a derivative. And, we're going to look at this from several different points of view, and the first one is the geometric interpretation. That's what we'll spend most of today on. And then, we'll also talk about a physical interpretation of what a derivative is.

And then there's going to be something else which I guess is maybe the reason why Calculus is so fundamental, and why we always start with it in most science and engineering schools, which is the importance of derivatives, of this, to all measurements. So that means pretty much every place. That means in science, in engineering, in economics, in political science, etc. Polling, lots of commercial applications, just about everything.

Now, that's what we'll be getting started with, and then there's another thing that we're gonna do in this unit, which is we're going to explain how to differentiate anything. So, how to differentiate any function you know. And that's kind of a tall order, but let me just give you an example. If you want to take the derivative - this we'll see today is the notation for the derivative of something - of some messy function like e ^ x arctan x. We'll work this out by the end of this unit.

All right? Anything you can think of, anything you can write down, we can differentiate it. All right, so that's what we're gonna do, and today, as I said, we're gonna spend most of our time on this geometric interpretation. So let's begin with that.

So here we go with the geometric interpretation of derivatives. And, what we're going to do is just ask the geometric problem of finding the tangent line to some graph of some function at some point. Which is to say (x_0, y_0) . So that's the problem that we're addressing here. Alright, so here's our problem, and now let me show you the solution. So, well, let's graph the function. Here's its graph. Here's some point. All right, maybe I should draw it just a bit lower. So here's a point P. Maybe it's above the point x_0 . x_0 , by the way, this was supposed to be

an x_0. That was some fixed place on the x-axis. And now, in order to perform this mighty feat, I will use another color of chalk. How about red? OK. So here it is. There's the tangent line, well, not quite straight. Close enough. All right? I did it.

That's the geometric problem. I achieved what I wanted to do, and it's kind of an interesting question, which unfortunately I can't solve for you in this class, which is, how did I do that? That is, how physically did I manage to know what to do to draw this tangent line? But that's what geometric problems are like. We visualize it. We can figure it out somewhere in our brains. It happens. And the task that we have now is to figure out how to do it analytically, to do it in a way that a machine could just as well as I did in drawing this tangent line.

So, what did we learn in high school about what a tangent line is? Well, a tangent line has an equation, and any line through a point has the equation $y - y_0$ is equal to m, the slope, times $x - x_0$.

So here's the equation for that line, and now there are two pieces of information that we're going to need to work out what the line is. The first one is the point. That's that point P there. And to specify P, given x, we need to know the level of y, which is of course just $f(x_0)$. That's not a calculus problem, but anyway that's a very important part of the process. So that's the first thing we need to know. And the second thing we need to know is the slope. And that's this number m. And in calculus we have another name for it. We call it f prime of x_0 . Namely, the derivative of f. So that's the calculus part. That's the tricky part, and that's the part that we have to discuss now. So just to make that explicit here, I'm going to make a definition, which is that $f'(x_0)$, which is known as the derivative, of f, at x_0 , is the slope of the tangent line to f'(x) at the point, let's just call it P.

All right? So, that's what it is, but still I haven't made any progress in figuring out any better how I drew that line. So I have to say something that's more concrete, because I want to be able to cook up what these numbers are. I have to figure out what this number m is. And one way of thinking about that, let me just try this, so I certainly am taking for granted that in sort of non-calculus part that I know what a line through a point is. So I know this equation. But another possibility might be, this line here, how do I know - well, unfortunately, I didn't draw it quite straight, but there it is - how do I know that this orange line is not a tangent line, but this other line is a tangent line? Well, it's actually not so obvious, but I'm gonna describe it a little bit. It's not really the fact-- this thing crosses at some other place, which is this point Q. But it's not really the fact that the thing crosses at two place, because the line could be wiggly, the

curve could be wiggly, and it could cross back and forth a number of times. That's not what distinguishes the tangent line.

So I'm gonna have to somehow grasp this, and I'll first do it in language. And it's the following idea: it's that if you take this orange line, which is called a secant line, and you think of the point Q as getting closer and closer to P, then the slope of that line will get closer and closer to the slope of the red line. And if we draw it close enough, then that's gonna be the correct line. So that's really what I did, sort of in my brain when I drew that first line.

And so that's the way I'm going to articulate it first. Now, so the tangent line is equal to the limit of so called secant lines PQ, as Q tends to P. And here we're thinking of P as being fixed and Q as variable. All right? Again, this is still the geometric discussion, but now we're gonna be able to put symbols and formulas to this computation. And we'll be able to work out formulas in any example.

So let's do that. So first of all, I'm gonna write out these points P and Q again. So maybe we'll put P here and Q here. And I'm thinking of this line through them. I guess it was orange, so we'll leave it as orange. All right. And now I want to compute its slope. So this, gradually, we'll do this in two steps. And these steps will introduce us to the basic notations which are used throughout calculus, including multi-variable calculus, across the board. So the first notation that's used is you imagine here's the x-axis underneath, and here's the x_0, the location directly below the point P. And we're traveling here a horizontal distance which is denoted by delta x. So that's delta x, so called. And we could also call it the change in x.

So that's one thing we want to measure in order to get the slope of this line PQ. And the other thing is this height. So that's this distance here, which we denote delta f, which is the change in f. And then, the slope is just the ratio, delta f / delta x. So this is the slope of the secant. And the process I just described over here with this limit applies not just to the whole line itself, but also in particular to its slope. And the way we write that is the limit as delta x goes to 0. And that's going to be our slope. So this is the slope of the tangent line.

OK. Now, This is still a little general, and I want to work out a more usable form here, a better formula for this. And in order to do that, I'm gonna write delta f, the numerator more explicitly here. The change in f, so remember that the point P is the point $(x_0, f(x_0))$. All right, that's what we got for the formula for the point. And in order to compute these distances and in particular the vertical distance here, I'm gonna have to get a formula for Q as well. So if this

horizontal distance is delta x, then this location is $x_0 + delta x$. And so the point above that point has a formula, which is $x_0 + delta x$. O plus delta x, f of - and this is a mouthful - $x_0 + delta x$.

All right, so there's the formula for the point Q. Here's the formula for the point P. And now I can write a different formula for the derivative, which is the following: so this $f'(x_0)$, which is the same as m, is going to be the limit as delta x goes to 0 of the change in f, well the change in f is the value of f at the upper point here, which is $x_0 + delta x$, and minus its value at the lower point P, which is $f(x_0)$, divided by delta x. All right, so this is the formula. I'm going to put this in a little box, because this is by far the most important formula today, which we use to derive pretty much everything else. And this is the way that we're going to be able to compute these numbers.

So let's do an example. This example, we'll call this example one. We'll take the function f(x), which is 1/x. That's sufficiently complicated to have an interesting answer, and sufficiently straightforward that we can compute the derivative fairly quickly. So what is it that we're gonna do here? All we're going to do is we're going to plug in this formula here for that function. That's all we're going to do, and visually what we're accomplishing is somehow to take the hyperbola, and take a point on the hyperbola, and figure out some tangent line. That's what we're accomplishing when we do that. So we're accomplishing this geometrically but we'll be doing it algebraically. So first, we consider this difference delta f / delta x and write out its formula.

So I have to have a place. So I'm gonna make it again above this point x_0 , which is the general point. We'll make the general calculation. So the value of f at the top, when we move to the right by f(x), so I just read off from this, read off from here. The formula, the first thing I get here is $1/(x_0 + \text{delta } x)$. That's the left hand term. Minus $1/x_0$, that's the right hand term. And then I have to divide that by delta x. OK, so here's our expression. And by the way this has a name. This thing is called a difference quotient. It's pretty complicated, because there's always a difference in the numerator. And in disguise, the denominator is a difference, because it's the difference between the value on the right side and the value on the left side here. OK, so now we're going to simplify it by some algebra.

So let's just take a look. So this is equal to, let's continue on the next level here. This is equal to 1 / delta x times... All I'm going to do is put it over a common denominator. So the common denominator is $(x_0 + \text{delta x}) * x_0$. And so in the numerator for the first expressions I have x_0 , and for the second expression I have $x_0 + \text{delta x}$. So this is the same thing as I had in

the numerator before, factoring out this denominator. And here I put that numerator into this more amenable form.

And now there are two basic cancellations. The first one is that x_0 and x_0 cancel, so we have this. And then the second step is that these two expressions cancel, the numerator and the denominator. Now we have a cancellation that we can make use of. So we'll write that under here. And this is equals -1 over x_0 plus delta x times x_0 . And then the very last step is to take the limit as delta x tends to 0, and now we can do it. Before we couldn't do it. Why? Because the numerator and the denominator gave us 0 / 0. But now that I've made this cancellation, I can pass to the limit. And all that happens is I set this delta x to 0, and I get - $1/x_0^2$. So that's the answer. All right, so in other words what I've shown - let me put it up here - is that $f'(x_0) = -1/x_0^2$.

Now, let's look at the graph just a little bit to check this for plausibility, all right? What's happening here is, first of all it's negative. It's less than 0, which is a good thing. You see that slope there is negative. That's the simplest check that you could make. And the second thing that I would just like to point out is that as x goes to infinity, that as we go farther to the right, it gets less and less steep. So as x_0 goes to infinity, less and less steep. So that's also consistent here, when x_0 is very large, this is a smaller and smaller number in magnitude, although it's always negative. It's always sloping down. All right, so I've managed to fill the boards. So maybe I should stop for a question or two. Yes?

Student:

[INAUDIBLE]

Professor:

So the question is to explain again this limiting process. So the formula here is we have basically two numbers. So in other words, why is it that this expression, when delta x tends to 0, is equal to $-1/x_0^2$? Let me illustrate it by sticking in a number for x_0 to make it more explicit. All right, so for instance, let me stick in here for x_0 the number 3. Then it's -1 over 3 plus delta x times 3. That's the situation that we've got. And now the question is what happens as this number gets smaller and smaller, and gets to be practically 0? Well, literally what we can do is just plug in 0 there, and you get 3 plus 0 times 3 in the denominator. -1 in the numerator. So this tends to -1/9 (over 3^2). And that's what I'm saying in general with this extra number here. Other questions? Yes.

Student:

[INAUDIBLE]

Professor:

So the question is what happened between this step and this step, right? Explain this step

here. Alright, so there were two parts to that. The first is this delta x which is sitting in the denominator, I factored all the way out front. And so what's in the parentheses is supposed to be the same as what's in the numerator of this other expression. And then, at the same time as doing that, I put that expression, which is the difference of two fractions, I expressed it with a common denominator. So in the denominator here, you see the product of the denominators of the two fractions. And then I just figured out what the numerator had to be without really... Other questions? OK.

So I claim that on the whole, calculus gets a bad rap, that it's actually easier than most things. But there's a perception that it's harder. And so I really have a duty to give you the calculus made harder story here. So we have to make things harder, because that's our job. And this is actually what most people do in calculus, and it's the reason why calculus has a bad reputation. So the secret is that when people ask problems in calculus, they generally ask them in context. And there are many, many other things going on. And so the little piece of the problem which is calculus is actually fairly routine and has to be isolated and gotten through. But all the rest of it, relies on everything else you learned in mathematics up to this stage, from grade school through high school. So that's the complication. So now we're going to do a little bit of calculus made hard. By talking about a word problem.

We only have one sort of word problem that we can pose, because all we've talked about is this geometry point of view. So far those are the only kinds of word problems we can pose. So what we're gonna do is just pose such a problem. So find the areas of triangles, enclosed by the axes and the tangent to y = 1/x. OK, so that's a geometry problem. And let me draw a picture of it. It's practically the same as the picture for example one. We only consider the first quadrant. Here's our shape. All right, it's the hyperbola. And here's maybe one of our tangent lines, which is coming in like this. And then we're trying to find this area here. Right, so there's our problem. So why does it have to do with calculus? It has to do with calculus because there's a tangent line in it, so we're gonna need to do some calculus to answer this question. But as you'll see, the calculus is the easy part.

So let's get started with this problem. First of all, I'm gonna label a few things. And one important thing to remember of course, is that the curve is y = 1/x. That's perfectly reasonable to do. And also, we're gonna calculate the areas of the triangles, and you could ask yourself, in terms of what? Well, we're gonna have to pick a point and give it a name. And since we need a number, we're gonna have to do more than geometry. We're gonna have to do some

of this analysis just as we've done before. So I'm gonna pick a point and, consistent with the labeling we've done before, I'm gonna to call it (x_0, y_0) . So that's almost half the battle, having notations, x and y for the variables, and x 0 and y 0, for the specific point.

Now, once you see that you have these labelings, I hope it's reasonable to do the following. So first of all, this is the point x_0, and over here is the point y_0. That's something that we're used to in graphs. And in order to figure out the area of this triangle, it's pretty clear that we should find the base, which is that we should find this location here. And we should find the height, so we need to find that value there. Let's go ahead and do it. So how are we going to do this? Well, so let's just take a look. So what is it that we need to do? I claim that there's only one calculus step, and I'm gonna put a star here for this tangent line. I have to understand what the tangent line is. Once I've figured out what the tangent line is, the rest of the problem is no longer calculus. It's just that slope that we need. So what's the formula for the tangent line? Put that over here. it's going to be y - y_0 is equal to, and here's the magic number, we already calculated it. It's in the box over there. It's -1/x_0^2 (x - x_0). So this is the only bit of calculus in this problem. But now we're not done. We have to finish it. We have to figure out all the rest of these quantities so we can figure out the area.

All right. So how do we do that? Well, to find this point, this has a name. We're gonna find the so called x-intercept. That's the first thing we're going to do. So to do that, what we need to do is to find where this horizontal line meets that diagonal line. And the equation for the x-intercept is y = 0. So we plug in y = 0, that's this horizontal line, and we find this point. So let's do that into star. We get 0 minus, oh one other thing we need to know. We know that y0 is $f(x_0)$, and f(x) is 1/x, so this thing is $1/x_0$. And that's equal to $-1/x_0^2$. And here's x, and here's x_0 . All right, so in order to find this x value, I have to plug in one equation into the other.

So this simplifies a bit. This is $-x/x_0^2$. And this is plus $1/x_0$ because the x_0 and $x0^2$ cancel somewhat. And so if I put this on the other side, I get x / x_0^2 is equal to $2 / x_0$. And if I then multiply through - so that's what this implies - and if I multiply through by x_0^2 I get $x = 2x_0$.

OK, so I claim that this point we've just calculated, it's 2x_0. Now, I'm almost done. I need to get the other one. I need to get this one up here. Now I'm gonna use a very big shortcut to do that. So the shortcut to the y-intercept is to use symmetry. All right, I claim I can stare at this and I can look at that, and I know the formula for the y-intercept. It's equal to 2y 0. All right.

That's what that one is. So this one is $2y_0$. And the reason I know this is the following: so here's the symmetry of the situation, which is not completely direct. It's a kind of mirror symmetry around the diagonal. It involves the exchange of (x, y) with (y, x); so trading the roles of x and y. So the symmetry that I'm using is that any formula I get that involves x's and y's, if I trade all the x's and replace them by y's and trade all the y's and replace them by x's, then I'll have a correct formula on the other way. So if everywhere I see a y I make it an x, and everywhere I see an x I make it a y, the switch will take place. So why is that? That's just an accident of this equation. That's because, so the symmetry explained... is that the equation is y = 1/x. But that's the same thing as x = 1/y. So here's where the x and the y get reversed. OK now if you don't trust this explanation, you can also get the y-intercept by plugging x = 0 into the equation star. OK? We plugged y = 0 in and we got the x-value. And you can do the same thing analogously the other way.

All right so I'm almost done with the geometry problem, and let's finish it off now. Well, let me hold off for one second before I finish it off. What I'd like to say is just make one more tiny remark. And this is the hardest part of calculus in my opinion. So the hardest part of calculus is that we call it one variable calculus, but we're perfectly happy to deal with four variables at a time or five, or any number. In this problem, I had an x, a y, an x_0 and a y_0. That's already four different things that have various relationships between them. Of course the manipulations we do with them are algebraic, and when we're doing the derivatives we just consider what's known as one variable calculus. But really there are millions of variable floating around potentially. So that's what makes things complicated, and that's something that you have to get used to. Now there's something else which is more subtle, and that I think many people who teach the subject or use the subject aren't aware, because they've already entered into the language and they're so comfortable with it that they don't even notice this confusion. There's something deliberately sloppy about the way we deal with these variables.

The reason is very simple. There are already four variables here. I don't wanna create six names for variables or eight names for variables. But really in this problem there were about eight. I just slipped them by you. So why is that? Well notice that the first time that I got a formula for y_0 here, it was this point. And so the formula for y_0 , which I plugged in right here, was from the equation of the curve. $y_0 = 1 / x_0$. The second time I did it, I did not use y = 1/x. I used this equation here, so this is not y = 1/x. That's the wrong thing to do. It's an easy mistake to make if the formulas are all a blur to you and you're not paying attention to

where they are on the diagram.

You see that x-intercept calculation there involved where this horizontal line met this diagonal line, and y = 0 represented this line here. So the sloppiness is that y means two different things. And we do this constantly because it's way, way more complicated not to do it. It's much more convenient for us to allow ourselves the flexibility to change the role that this letter plays in the middle of a computation. And similarly, later on, if I had done this by this more straightforward method, for the y-intercept, I would have set x equal to 0. That would have been this vertical line, which is x = 0. But I didn't change the letter x when I did that, because that would be a waste for us. So this is one of the main confusions that happens. If you can keep yourself straight, you're a lot better off, and as I say this is one of the complexities.

All right, so now let's finish off the problem. Let me finally get this area here. So, actually I'll just finish it off right here. So the area of the triangle is, well it's the base times the height. The base is $2x_0$, the height is $2y_0$, and a half of that. So it's 1/2 $(2x_0)$ * $(2y_0)$, which is $2x_0$ y_0 , which is, lo and behold, 2. So the amusing thing in this case is that it actually didn't matter what x_0 and y_0 are. We get the same answer every time. That's just an accident of the function 1/x. It happens to be the function with that property.

All right, so we have some more business today, some serious business. So let me continue. So, first of all, I want to give you a few more notations. And these are just other notations that people use to refer to derivatives. And the first one is the following: we already wrote y = f(x). And so when we write delta y, that means the same thing as delta f. That's a typical notation. And previously we wrote f prime for the derivative, so this is Newton's notation for the derivative. But there are other notations. And one of them is df/dx, and another one is dy/dx, meaning exactly the same thing. And sometimes we let the function slip down below so that becomes d/dx of f and d/dx of y. So these are all notations that are used for the derivative, and these were initiated by Leibniz. And these notations are used interchangeably, sometimes practically together. They both turn out to be extremely useful. This one omits - notice that this thing omits - the underlying base point, x_0. That's one of the nuisances. It doesn't give you all the information. But there are lots of situations like that where people leave out some of the important information, and you have to fill it in from context. So that's another couple of notations.

So now I have one more calculation for you today. I carried out this calculation of the derivative of the function 1 / x. I wanna take care of some other powers. So let's do that.

So Example 2 is going to be the function $f(x) = x^n$. n = 1, 2, 3; one of these guys. And now what we're trying to figure out is the derivative with respect to x of x^n in our new notation, what this is equal to. So again, we're going to form this expression, delta f / delta x. And we're going to make some algebraic simplification. So what we plug in for delta f is ((x delta x)^n x^n)/delta x. Now before, let me just stick this in then I'm gonna erase it. Before, I wrote x 0 here and x 0 there. But now I'm going to get rid of it, because in this particular calculation, it's a nuisance. I don't have an x floating around, which means something different from the x 0. And I just don't wanna have to keep on writing all those symbols. It's a waste of blackboard energy. There's a total amount of energy, and I've already filled up so many blackboards that, there's just a limited amount. Plus, I'm trying to conserve chalk. Anyway, no 0's. So think of x as fixed. In this case, delta x moves and x is fixed in this calculation. All right now, in order to simplify this, in order to understand algebraically what's going on, I need to understand what the nth power of a sum is. And that's a famous formula. We only need a little tiny bit of it, called the binomial theorem. So, the binomial theorem which is in your text and explained in an appendix, says that if you take the sum of two guys and you take them to the nth power, that of course is (x + delta x) multiplied by itself n times. And so the first term is x^n , that's when all of the n factors come in. And then, you could have this factor of delta x and all the rest x's. So at least one term of the form $(x^{(n-1)})$ times delta x. And how many times does that happen? Well, it happens when there's a factor from here, from the next factor, and so on, and so on, and so on. There's a total of n possible times that that happens. And now the great thing is that, with this alone, all the rest of the terms are junk that we won't have to worry about. So to be more specific, there's a very careful notation for the junk. The junk is what's called big O of (delta x)^2. What that means is that these are terms of order, so with (delta x)^2, (delta x)^3 or higher. All right, that's how. Very exciting, higher order terms. OK, so this is the only algebra that we need to do, and now we just need to combine it together to get our result. So, now I'm going to just carry out the cancellations that we need. So here we go. We have delta f / delta x, which remember was 1 / delta x times this, which is this times, now this is x^n plus nx^(n-1) delta x plus this junk term, minus x^n. So that's what we have so far based on our previous calculations. Now, I'm going to do the main cancellation, which is this. All right. So, that's 1/delta x times nx^(n-1) delta x plus this term here. And now I can divide in by delta x. So I get nx^(n-1) plus, now it's O(delta x). There's at least one factor of delta x not two factors of delta x, because I have to cancel one of them. And now I can just take the limit. In the limit this term is gonna be 0. That's why I called it junk originally, because it disappears. And in math, junk is something that goes away. So this tends to, as delta x goes to 0, $nx^{(n-1)}$. And so what I've

shown you is that d/dx of x to the n minus-- sorry, n, is equal to $nx^{(n-1)}$.

So now this is gonna be super important to you right on your problem set in every possible way, and I want to tell you one thing, one way in which it's very important. One way that extends it immediately. So this thing extends to polynomials. We get quite a lot out of this one calculation. Namely, if I take d/dx of something like $(x^3 + 5x^10)$ that's gonna be equal to $3x^2$, that's applying this rule to x^3 . And then here, I'll get 5^*10 so $50x^9$. So this is the type of thing that we get out of it, and we're gonna make more hay with that next time. Question. Yes. I turned myself off. Yes?

Student:

[INAUDIBLE]

Professor:

The question was the binomial theorem only works when delta x goes to 0. No, the binomial theorem is a general formula which also specifies exactly what the junk is. It's very much more detailed. But we only needed this part. We didn't care what all these crazy terms were. It's junk for our purposes now, because we don't happen to need any more than those first two terms. Yes, because delta x goes to 0. OK, see you next time.