

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](http://ocw.mit.edu).

**Professor:** I am Haynes Miller, I am substituting for David Jerison today. So you have a substitute teacher today. So I haven't been here in this class with you so I'm not completely sure where you are. I think you've just been talking about differentiation and you've got some examples of differentiation like these basic examples: the derivative of  $x^n$  is  $nx^{(n-1)}$ . But I think maybe you've spent some time computing the derivative of the sine function as well, recently. And I think you have some rules for extending these calculations as well. For instance, I think you know that if you differentiate a constant times a function, what do you get?

**Student:** [INAUDIBLE].

**Professor:** The constant comes outside like this. Or I could write  $(cu)' = cu'$ . That's this rule, multiplying by a constant, and I think you also know about differentiating a sum. Or I could write this as  $(u + v)' = u' + v'$ . So I'm going to be using those but today I'll talk about a collection of other rules about how to deal with a product of functions, a quotient of functions, and, best of all, composition of functions. And then at the end, I'll have something to say about higher derivatives. So that's the story for today. That's the program.

So let's begin by talking about the product rule. So the product rule tells you how to differentiate a product of functions, and I'll just give you the rule, first of all. The rule is it's  $u'v + uv'$ . It's a little bit funny. Differentiating a product gives you a sum. But let's see how that works out in a particular example. For example, suppose that I wanted to differentiate the product. Well, the product of these two basic examples that we just talked about. I'm going to use the same variable in both cases instead of different ones like I did here. So the derivative of  $x^n$  times  $\sin x$ . So this is a new thing. We couldn't do this without using the product rule. So the first function is  $x^n$  and the second one is  $\sin x$ . And we're going to apply this rule. So  $u$  is  $x^n$ .  $u'$  is, according to the rule,  $nx^{(n-1)}$ . And then I take  $v$  and write it down the way it is, sine of  $x$ . And then I do it the other way. I take  $u$  the way it is, that's  $x^n$ , and multiply it by the derivative of  $v$ ,  $v'$ . We just saw  $v'$  is cosine of  $x$ . So that's it. Obviously, you can differentiate longer products, products of more things by doing it one at a time.

Let's see why this is true. I want to try to show you why the product rule holds. So you have a standard way of trying to understand this, and it involves looking at the change in the function that you're interested in differentiating. So I should look at how much the product  $uv$  changes when  $x$  changes a little bit. Well, so how do compute the change? Well, I write down the value of the function at some new value of  $x$ ,  $x + \Delta x$ . Well, I better write down the whole new value of the function, and the function is  $uv$ . So the whole new value looks like this. It's  $u(x + \Delta x)$  times  $v(x + \Delta x)$ . That's the new value. But what's the change in the product? Well, I better subtract off what the old value was, which is  $u(x)v(x)$ .

Okay, according to the rule we're trying to prove, I have to get  $u'$  involved. So I want to involve the change in  $u$  alone, by itself. Let's just try that. I see part of the formula for the change in  $u$  right there. Let's see if we can get the rest of it in place. So the change in  $u$  is  $(u(x + \Delta x) - u(x))$ . That's the change in  $u$  [Correction: change in  $u$ ]. This part of it occurs up here, multiplied by  $v(x + \Delta x)$ , so let's put that in too. Now this equality sign isn't very good right now. I've got this product here so far, but I've introduced something I don't like. I've introduced  $u$  times  $v(x + \Delta x)$ , right? Minus that. So the next thing I'm gonna do is correct that little defect by adding in  $u(x)v(x + \Delta x)$ . Okay, now I cancelled off what was wrong with this line. But I'm still not quite there, because I haven't put this in yet. So I better subtract off  $uv$ , and then I'll be home. But I'm going to do that in a clever way, because I noticed that I already have a  $u$  here. So I'm gonna take this factor of  $u$  and make it the same as this factor. So I get  $u(x)$  times this, minus  $u(x)$  times that. That's the same thing as  $u$  times the difference. So that was a little bit strange, but when you stand back and look at it, you can see multiplied out, the middle terms cancel. And you get the right answer. Well I like that because it's involved the change in  $u$  and the change in  $v$ . So this is equal to  $\Delta u$  times  $v(x + \Delta x)$  minus  $u(x)$  times the change in  $v$ . Well, I'm almost there. The next step in computing the derivative is take difference quotient, divide this by  $\Delta x$ . So,  $(\Delta(uv)) / (\Delta x)$  is well, I'll say  $(\Delta u / \Delta x)$  times  $v(x + \Delta x)$ .

Have I made a mistake here? This plus magically became a minus on the way down here, so I better fix that. Plus  $u$  times  $(\Delta v) / (\Delta x)$ . This  $u$  is this  $u$  over here. So I've just divided this formula by  $\Delta x$ , and now I can take the limit as  $x$  goes to 0, so this is as  $\Delta x$  goes to 0. This becomes the definition of the derivative, and on this side, I get  $du/dx$  times... now what happens to this quantity when  $\Delta x$  goes to 0? So this quantity is getting closer and closer to  $x$ . So what happens to the value of  $v$ ? It becomes equal to  $x$  of  $v$ . That uses continuity of  $v$ . So,  $v(x + \Delta x)$  goes to  $v(x)$  by continuity. So this gives me times  $v$ , and then I have  $u$  times, and

$\Delta v / \Delta x$  gives me  $dv/dx$ . And that's the formula. That's the formula as I wrote it down at the beginning over here. The derivative of a product is given by this sum. Yeah?

**Student:** How did you get from the first line to the second of the long equation?

**Professor:** From here to here?

**Student:** Yes.

**Professor:** So maybe it's easiest to work backwards and verify that what I wrote down is correct here. So, if you look there's a  $u$  times  $v(x + \Delta x)$  there. And there's also one here. And they occur with opposite signs. So they cancel. What's left is  $u(x + \Delta x) - uv$ . And that's just what I started with.

**Student:** [INAUDIBLE] They cancel right?

**Professor:** I cancelled out this term and this term, and what's left is the ends. Any other questions?

**Student:** [INAUDIBLE].

**Professor:** Well, I just calculated what  $\Delta uv$  is, and now I'm gonna divide that by  $\Delta x$  on my way to computing the derivative. And so I copied down the right hand side and divided  $\Delta x$ . I just decided to divide the  $\Delta u$  by  $\Delta x$  and  $\Delta v$  by  $\Delta x$ . Good. Anything else?

So we have the product rule here. The rule for differentiating a product of two functions. This is making us stronger. There are many more functions you can find derivatives of now. How about quotients? Let's find out how to differentiate a quotient of two functions. Well again, I'll write down what the answer is and then we'll try to verify it. So there's a quotient. Let me write this down. There's a quotient of two functions. And here's the rule for it. I always have to think about this and hope that I get it right.  $(u'v - uv') / v^2$ . This may be the craziest rule you'll see in this course, but there it is. And I'll try to show you why that's true and see an example. Yeah there was a hand?

**Student:** [INAUDIBLE]

**Professor:** What letters look the same?  $u$  and  $v$  look the same? I'll try to make them look more different. The  $v$ 's have points on the bottom.  $u$ 's have little round things on the bottom. What's the new value of  $u$ ? The value of  $u$  at  $x + \Delta x$  is  $u + \Delta u$ , right? That's what  $\Delta u$  is. It's the change in  $u$  when  $x$  gets replaced by  $\Delta x$  [Correction:  $x + \Delta x$ ]. And the change in  $v$ , the

new value  $v$ , is  $v + \Delta v$ . So this is the new value of  $u$  divided by the new value of  $v$ . That's the beginning. And then I subtract off the old values, which are  $u$  minus  $v$ . This'll be easier to work out when I write it out this way. So now, we'll cross multiply, as I said. So I get  $uv + (\Delta u)v$  minus, now I cross multiply this way, you get  $uv - u(\Delta v)$ . And I divide all this by  $(v + \Delta v)u$ . Okay, now the reason I like to do it this way is that you see the cancellation happening here.  $uv$  and  $uv$  occur twice and so I can cancel them. And I will, and I'll answer these questions in a minute.

**Audience:** [INAUDIBLE].

**Professor:** Ooh, that's a  $v$ . All right. Good, anything else? That's what all hands were. Good. All right, so I cancel these and what I'm left with then is  $\Delta u$  times  $v$  minus  $u$  times  $\Delta v$  and all this is over  $v + \Delta v$  times  $v$ . Okay, there's the difference. There's the change in the quotient. The change in this function is given by this formula. And now to compute the derivative, I want to divide by  $\Delta x$ , and take the limit. So let's write that down,  $\Delta(u/v)/\Delta x$  is this formula here divided by  $\Delta x$ . And again, I'm going to put the  $\Delta x$  under these  $\Delta u$  and  $\Delta v$ . Okay? I'm gonna put  $\Delta x$  in the denominator, but I can think of that as dividing into this factor and this factor. So this is  $(\Delta u / \Delta x)v - u(\Delta v / \Delta x)$ . And all that is divided by the same denominator,  $(v + \Delta v)v$ . Right? Put the  $\Delta x$  up in the numerator there.

Next up, take the limit as  $\Delta x$  goes to 0. I get, by definition, the derivative of  $(u/v)$ . And on the right hand side, well, this is the derivative  $du/dx$  right? Times  $v$ . See and then  $u$  times, and here it's the derivative  $dv/dx$ . Now what about the denominator? So when  $\Delta x$  goes to 0,  $v$  stays the same,  $v$  stays the same. What happens to this  $\Delta v$ ? It goes to 0, again, because  $v$  is continuous. So again,  $\Delta v$  goes to 0 with  $\Delta x$  because they're continuous and you just get  $v$  times  $v$ . I think that's the formula I wrote down over there.  $(du/dx)v - u(dv/dx)$ . And all divided by the square of the old denominator. Well, that's it. That's the quotient rule. Weird formula.

Let's see an application. Let's see an example. So the example I'm going to give is pretty simple. I'm going to take the numerator to be just 1. So I'm gonna take  $u = 1$ . So now I'm differentiating  $1/v$ , the reciprocal of a function; 1 over a function. Here's a copy of my rule. What's  $du/dx$  in that case?  $u$  is a constant, so that term is 0 in this rule. I don't have to worry about this. I get a minus. And then  $u$  is 1, and  $dv/dx$ . Well,  $v$  is whatever  $v$  is. I'll write  $dv/dx$  as  $v'$ . And then I get a  $v^2$  in the denominator. So that's the rule. I could write it as  $v^{-2} v'$ . Minus  $v'$  divided by  $v^2$ . That's the derivative of  $1/v$ .

How about sub-example of that? I'm going to take the special case where  $u = 1$  again. And  $v$  is a power of  $x$ . And I'm gonna use the rule that we developed earlier about the derivative of  $x^n$ . So what do I get here?  $d/dx (1/x^n)$  is, I'm plugging into this formula here with  $v = x^n$ . So I get minus, uh,  $v^{-2}$ . If  $v = x^n$ ,  $v^{-2}$  is, by the rule of exponents,  $x^{-2n}$ . And then  $v'$  is the derivative of  $x^n$ , which is  $nx^{n-1}$ .

Okay, so let's put these together. There's several powers of  $x$  here. I can put them together. I get  $-n x$  to the  $-2n + n - 1$ . One of these  $n$ 's cancels. And what I'm left with is  $-n - 1$ . So we've computed the derivative of  $1/x^n$ , which I could also write as  $x^{-n}$ , right? So I've computed the derivative of negative powers of  $x$ . And this is the formula that I get. If you think of this  $-n$  as a unit, as a thing to itself, it occurs here in the exponent. It occurs here, and it occurs here. So how does that compare with the formula that we had up here? The derivative of a power of  $x$  is that power times  $x$  to one less than that power. That's exactly the same as the rule that I wrote down here. But the power here happens to be a negative number, and the same negative number shows up as a coefficient and there in the exponent. Yeah?

**Student:** [INAUDIBLE].

**Professor:** How did I do this?

**Student:** [INAUDIBLE].

**Professor:** Where did that  $x^{-2n}$  come from? So I'm applying this rule. So the denominator in the quotient rule is  $v^2$ . And  $v$  was  $x^n$ , so the denominator is  $x^{2n}$ . And I decided to write it as  $x^{-2n}$ . So the green comments there... What they say is that I can enlarge this rule. This exact same rule is true for negative values of  $n$ , as well as positive values of  $n$ . So there's something new in your list of rules that you can apply, of values of the derivative. That standard rule is true for negative as well as positive exponents. And that comes out of a quotient rule.

Okay, so we've done two rules. I've talked about the product rule and the quotient rule. What's next? Let's see the chain rule. So this is a composition rule. So the kind of thing that I have in mind, composition of functions is about substitution. So the kind of function that I have in mind is, for instance,  $y = (\sin t)^{10}$ . That's a new one. We haven't seen how to differentiate that before, I think. This kind of power of a trig function happens very often. You've seen them happen, as well, I'm sure, already. And there's a little notational switch that people use. They'll

write  $\sin^{10}(t)$ . But remember that when you write  $\sin^{10}(t)$ , what you mean is take the sine of  $t$ , and then take the 10th power of that. It's the meaning of  $\sin^{10}(t)$ . So the method of dealing with this kind of composition of functions is to use new variable names. What I mean is, I can think of this  $(\sin t)^{10}$ .

I can think of it as a two step process. First of all, I compute the sine of  $t$ . And let's call the result  $x$ . There's the new variable name. And then, I express  $y$  in terms of  $x$ . So  $y$  says take this and raise it to the tenth power. In other words,  $y = x^{10}$ . And then you plug  $x = \sin(t)$  into that, and you get the formula for what  $y$  is in terms of  $t$ . So it's good practice to introduce new letters when they're convenient, and this is one place where it's very convenient.

So let's find a rule for differentiating a composition, a function that can be expressed by doing one function and then applying another function. And here's the rule. Well, maybe I'll actually derive this rule first, and then you'll see what it is. In fact, the rule is very simple to derive. So this is a proof first, and then we'll write down the rule. I'm interested in  $\Delta y / \Delta t$ .  $y$  is a function of  $x$ .  $x$  is a function of  $t$ . And I'm interested in how  $y$  changes with respect to  $t$ , with respect to the original variable  $t$ . Well, because of that intermediate variable, I can write this as  $(\Delta y / \Delta x) (\Delta x / \Delta t)$ . It cancels, right? The  $\Delta x$  cancels. The change in that immediate variable cancels out. This is just basic algebra. But what happens when I let  $\Delta t$  get small? Well this give me  $dy/dt$ . On the right-hand side, I get  $(dy/dx) (dx/dt)$ . So students will often remember this rule. This is the rule, by saying that you can cancel out for the  $dx$ 's. And that's not so far from the truth. That's a good way to think of it.

In other words, this is the so-called chain rule. And it says that differentiation of a composition is a product. It's just the product of the two derivatives. So that's how you differentiate a composite of two functions. And let's just do an example. Let's do this example. Let's see how that comes out. So let's differentiate, what did I say?  $(\sin t)^{10}$ . Okay, there's an inside function and an outside function. The inside function is  $x$  as a function of  $t$ . This is the inside function, and this is the outside function. So the rule says, first of all let's differentiate the outside function. Take  $dy/dx$ . Differentiate it with respect to that variable  $x$ . The outside function is the 10th power. What's its derivative? So I get  $10x^9$ . In this account, I'm using this newly introduced variable named  $x$ . So the derivative of the outside function is  $10x^9$ . And then here's the inside function, and the next thing I want to do is differentiate it. So what's  $dx/dt$ ,  $d/dt (\sin t)$ , the derivative of sine  $t$ ? All right, that's cosine  $t$ . That's what the chain rule gives you. This is correct, but since we were the ones to introduce this notation  $x$  here, that

wasn't given to us in the original problem here. The last step in this process should be to put back, to substitute back in what  $x$  is in terms of  $t$ . So  $x = \sin t$ . So that tells me that I get  $10(\sin(t))^9$ , that's  $x^9$ , times the  $\cos(t)$ . Or the same thing is  $\sin^9(t)\cos(t)$ . So there's an application of the chain rule.

You know, people often wonder where the name chain rule comes from. I was just wondering about that myself. So is it because it chains you down? Is it like a chain fence? I decided what it is. It's because by using it, you burst the chains of differentiation, and you can differentiate many more functions using it. So when you want to think of the chain rule, just think of that chain there. It lets you burst free.

Let me give you another application of the chain rule. Ready for this one? So I'd like to differentiate the  $\sin(10t)$ . Again, this is the composite of two functions. What's the inside function? Okay, so I think I'll introduce this new notation.  $x = 10t$ , and the outside function is the sine. So  $y = \sin x$ . So now the chain rule says  $dy/dt$  is... Okay, let's see. I take the derivative of the outside function, and what's that? Sine prime and we can substitute because we know what sine prime is. So I get cosine of whatever,  $x$ , and then times what? Now I differentiate the inside function, which is just 10. So I could write this as  $10\cos$  of what?  $10t$ ,  $x$  is  $10t$ . Now, once you get used to this, this middle variable, you don't have to give a name for it. You can just think about it in your mind without actually writing it down,  $d/dt(\sin(10t))$ . I'll just do it again without introducing this middle variable explicitly. Think about it. I first of all differentiate the outside function, and I get cosine. But I don't change the thing that I'm plugging into it. It's still  $x$  that I'm plugging into it.  $x$  is  $10t$ . So let's just write  $10t$  and not worry about the name of that extra variable. If it confuses you, introduce the new variable. And do it carefully and slowly like this. But, quite quickly, I think you'll get to be able to keep that step in your mind. I'm not quite done yet. I haven't differentiated the inside function, the derivative of  $10t = 10$ . So you get, again, the same result. A little shortcut that you'll get used to. Really and truly, once you have the chain rule, the world is yours to conquer. It puts you in a very, very powerful position.

Okay, well let's see. What have I covered today? I've talked about product rule, quotient rule, composition. I should tell you something about higher derivatives, as well. So let's do that. This is a simple story. Higher is kind of a strange word. It just means differentiate over and over again. All right, so let's see. If we have a function  $u$  or  $u(x)$ , please allow me to just write it as briefly as  $u$ . Well, this is a sort of notational thing. I can differentiate it and get  $u'$ . That's a new

function. Like if you started with the sine, that's gonna be the cosine. A new function, so I can differentiate it again. And the notation for the differentiating of it again, is  $u''$ . So  $u''$  is just  $u'$  differentiated again. For example, if  $u$  is the sine of  $x$ , so  $u'$  is the cosine of  $x$ . Has Professor Jerison talked about what the derivative of cosine is? What is it? Ha, okay so  $u''$  is  $-\sin x$ . Let me go on. What do you suppose  $u'''$  means? I guess it's the derivative of  $u''$ . It's called the third derivative. And  $u''$  is called the second derivative. And it's  $u''$  differentiated again. So to compute  $u'''$  in this example, what do I do? I differentiate that again. There's a constant term,  $-1$ , constant factor. That comes out. The derivative of sine is what? Okay, so  $u''' = -\cos x$ . Let's do it again. Now after a while, you get tired of writing these things. And so maybe I'll use the notation  $u^{(4)}$ . That's the fourth derivative. That's  $u^{(4)}$ . Or it's  $(u''')$  differentiated again, the fourth derivative. And what is that in this example? Okay, the cosine has derivative minus the sine, like you told me. And that minus sign cancels with that sign, and all together, I get  $\sin x$ . That's pretty bizarre. When I differentiate the function sine of  $x$  four times, I get back to the sine of  $x$  again. That's the way it is.

Now this notation, prime prime prime prime, and things like that. There are different variants of that notation. For example, that's another notation. Well, you've used the notation  $du/dx$  before.  $u'$  could also be denoted  $du/dx$ . I think we've already here, today, used this way of rewriting  $du/dx$ . I think when I was talking about  $d/dt(uv)$  and so on, I pulled that  $d/dt$  outside and put whatever function you're differentiating over to the right. So that's just a notational switch. It looks good. It looks like good algebra doesn't it? But what it's doing is regarding this notation as an operator. It's something you apply to a function to get a new function. I apply it to the sine function, and I get the cosine function. I apply it to  $x^2$ , and I get  $2x$ . This thing here, that symbol, represents an operator, which you apply to a function. And the operator says, take the function and differentiate it. So further notation that people often use, is they give a different name to that operator. And they'll write capital  $D$  for it. So this is just using capital  $D$  for the symbol  $d/dx$ .

So in terms of that notation, let's see. Let's write down what higher derivatives look like. So let's see. That's what  $u'$  is. How about  $u''$ ? Let's write that in terms of the  $d/dx$  notation. Well I'm supposed to differentiate  $u'$  right? So that's  $d/dx$  applied to the function  $du/dx$ . Differentiate the derivative. That's what I've done. Or I could write that as  $d/dx$  applied to  $d/dx$  applied to  $u$ . Just pulling that  $u$  outside. So I'm doing  $d/dx$  twice. I'm doing that operator twice. I could write that as  $(d/dx)^2$  applied to  $u$ . Differentiate twice, and do it to the function  $u$ . Or, I can write it as, now this is a strange one. I could also write as-- like that. It's getting stranger and stranger,



isn't it? This is definitely just a kind of abuse of notation. But people will go even further and write  $d^2 u / dx^2$ . So this is the strangest one. This identity quality is the strangest one, because you may think that you're taking  $d$  of the quantity  $x^2$ . But that's not what's intended. This is not  $d(x^2)$ . What's intended is the quantity  $dx^2$ . In this notation, which is very common, what's intended by the denominator is the quantity  $dx^2$ . It's part of this second differentiation operator. So I've written a bunch of equalities down here, and the only content to them is that these are all different notations for the same thing. You'll see this notation very commonly. So for instance the third derivative is  $d^3 u / dx^3$ , and so on. Sorry?

**Student:** [INAUDIBLE]. Professor: Yes, absolutely. Or an equally good notation is to write the operator capital  $D$ , done three times, to  $u$ . Absolutely. So I guess I should also write over here when I was talking about  $d^2$ , the second derivative, another notation is do the operator capital  $D$  twice. Let's see an example of how this can be applied. I'll answer this question.

**Student:** [INAUDIBLE].

**Professor:** Yeah, so the question is whether the fourth derivative always gives you the original function back, like what happened here. No. That's very, very special to sines and cosines. All right? And, in fact, let's see an example of that. I'll do a calculation. Let's calculate the  $n$ th derivative of  $x^n$ . Okay,  $n$  is a number, like 1, 2, 3, 4. Here we go. Let's do this. So, let's do this bit by bit. What's the first derivative of  $x^n$ ? So everybody knows this. I'm just using a new notation, this capital  $D$  notation. So it's  $nx^{n-1}$ . Now know, by the way,  $n$  could be a negative number for that, but for now, for this application, I wanna take  $n$  to be 1, 2, 3, and so on; one of those numbers.

Okay, we did one derivative. Let's compute the second derivative of  $x^n$ . Well there's this  $n$  constant that comes out, and then the exponent comes down, and it gets reduced by 1. All right? Should I do one more?  $D^3(x^n)$  is  $n(n-1)$ . That's the constant from here. Times that exponent,  $n-2$ , times 1 less,  $n-3$  is the new exponent. Well, I keep on going until I come to a new blackboard. Now, I think I'm going to stop when I get to the  $n$  minus first derivative, so we can see what's likely to happen. So when I took the third derivative, I had the  $n$  minus third power of  $x$ . And when I took the second derivative, I had the second power of  $x$ . So, I think what'll happen when I have the  $n$  minus first derivative is I'll have the first power of  $x$  left over. The powers of  $x$  keep coming down. And what I've done it  $n-1$  times, I get the first power. And then I get a big constant out in front here times more and more and more of these smaller

and smaller integers that come down. What's the last integer that came down before I got  $x^1$  here? Well, let's see. It's just 2, because this  $x^1$  occurred as the derivative of  $x^2$ . And the coefficient in front of that is 2. So that's what you get. The numbers  $n$ ,  $n-1$ , and so on down to 2, times  $x^1$ . And now we can differentiate one more time and calculate what  $D^n x^n$  is. So I get the same number,  $n$  times  $n-1$  and so on and so on, times 2. And then I guess I'll say times 1. Times, what's the derivative of  $x^1 = 1$ , so times 1. Time 1, times 1. Where this one means the constant function 1. Does anyone know what this number is called? That has a name. It's called  $n$  factorial. And it's written  $n$  exclamation point. And we just used an example of mathematical induction. So the end result is  $D^n x^n$  is  $n!$ , constant. Okay that's a neat fact. Final question for the lecture is what's  $D^{(n+1)}$  applied to  $x^n$ ? Ha. Excellent. It's the derivative of a constant. So it's 0.

Okay. Thank you.