PROFESSOR: Hi. Well, today's the chain rule. Very, very useful rule, and it's kind of neat, natural. Can I explain what a chain of functions is? There is a chain of functions. And then we want to know the slope, the derivative.

So how does the chain work? So there x is the input. It goes into a function g of x. We could call that inside function y. So the first step is y is g of x. So we get an output from g, call it y. That's halfway, because that y then becomes the input to f. That completes the chain. It starts with x, produces y, which is the inside function g of x. And then let me call it z is f of y.

And what I want to know is how quickly does z change as x changes? That's what the chain rule asks. It's the slope of that chain. Can I maybe just tell you the chain rule? And then we'll try it on some examples. You'll see how it works.

OK, here it is. The derivative, the slope of this chain dz dx, notice I want the change in the whole thing when I change the original input. Then the formula is that I take-- it's nice. You take dz dy times dy dx. So the derivative that we're looking for, the slope, the speed, is a product of two simpler derivatives that we probably know. And when we put the chain together, we multiply those derivatives. But there's one catch that I'll explain. I can give you a hint right away.

dz dy, this first factor, depends on y. But we're looking for the change due to the original change in x. When I find dz dy, I'm going to have to get back to x. Let me just do an example with a picture. You'll see why I have to do it.

So let the chain be cosine of-- oh, sine. Why not? Sine of 3x. Let me take sine of 3x. So that's my sine of 3x. I would like to know if that's my function, and I can draw it and will draw it, what is the slope?

OK, so what's the inside function? What's y here? Well, it's sitting there in parentheses. Often it's in parentheses so we identify it right away. y is 3x. That's the inside function. And then the outside function is the sine of y.

So what's the derivative by the chain rule? I'm ready to use the chain rule, because these are such simple functions, I know their separate derivative. So if this whole thing is z, the chain rule says that dz dx is-- I'm using this rule. I first name dz dy, the derivative of z with respect to y, which is cosine of y. And then the second factor is dy dx, and that's just a straight line with slope 3, so dy dx is 3. Good. Good, but not finished, because I'm getting an answer that's still in terms of y, and I have to get back to x, and no problem to do it. I know the link from y to x.

So here's the 3. I can usually write it out here, and then I wouldn't need parentheses. That's just that 3. Now the part I'm caring about: cosine of 3x. Not cosine x, even though this was just sine. But it was sine of y, and therefore, we need cosine 3x. Let me draw a picture of this function, and you'll see what's going on.

If I draw a picture of-- I'll start with a picture of sine x, maybe out to 180 degrees pi. This direction is now x. And this direction is going to be-- well, there is the sine of x, but that's not my function. My function is sine of 3x, and it's worth realizing what's the difference.

How does the graph change if I have 3x instead of x? Well, things come sooner. Things are speeded up. Here at x equal pi, 180 degrees, is when the sine goes back to 0. But for 3x, it'll be back to 0 already when x is 60 degrees, pi over 3. So 1/3 of the way along, right there, my sine 3x is this one. It's just like the sine curve but faster. That was pi over 3 there, 60 degrees.

So this is my z of x curve, and you can see that the slope is steeper at the beginning. You can see that the slope-- things are happening three times faster. Things are compressed by 3. This sine curve is compressed by 3. That makes it speed up so the slope is 3 at the start, and I claim that it's 3 cosine of 3x. Oh, let's draw the slope. All right, draw the slope.

All right, let me start with the slope of sine-- so this was just old sine x. So its slope is just cosine x along to-- right? That's the slope starts at 1. This is now cosine x. But that's going out to pi again. That's the slope of the original one, not the slope of our function, of our chain.

So the slope of our chain will be-- I mean, it doesn't go out so far. It's all between here and pi over 3, right? Our function, the one we're looking at, is just on this part. And the slope starts out at 3, and it's three times bigger, so it's going to be-- well, I'll just about get it on there. It's going to go down. I don't know if that's great, but it maybe makes the point that I started up here at 3, and I ended down here at minus 3 when x was 60 degrees because then-- you see, this is a picture of 3 cosine of 3x. I had to replace y by 3x at this point.

OK, let me do two or three more examples, just so you see it. Let's take an easy one. Suppose z is x cubed squared. All right, here is the inside function. y is x cubed, and z is-- do you see what z is? z is x cubed squared. So x cubed is the inside function.

What's the outside function? It's a function of y. I'm not going to write-- it's going to be the squaring function. That's what we do outside. I'm not going to write x squared. It's y squared. This is y. It's y squared that gets squared. Then the derivative dz dx by the chain rule is dz dy. Shall I remember the chain rule? dz dy, dy dx. Easy to remember because in the mind of everybody, these dy's, you see that they're sort of canceling.

So what's dz dy? z is y squared, so this is 2y, that factor. What's dy dx? y is x cubed. We know the derivative of x cubed. It's 3x squared.

There is the answer, but it's not final because I've got a y here that doesn't belong. I've got to get it back to. X So I

have all together 2 times 3 is making 6, and that y, I have to go back and see what was y in terms of x. It was x cubed. So I have x cubed there, and here's an x squared, altogether x to the fifth.

Now, is that the right answer? In this example, we can certainly check it because we know what x cubed squared is. So x cubed is x times x times x, and I'm squaring that. I'm multiplying by itself. There's another x times x times x. Altogether I have x to the sixth power. Notice I don't add those. When I'm squaring x cubed, I multiply the 2 by the 3 and get 6. So z is x to the sixth, and of course, the derivative of x to the sixth is 6 times x to the fifth, one power lower.

OK, I want to do two more examples. Let me do one more right away while we're on a roll. I'll bring down that board and take this function, just so you can spot the inside function and the outside function. So my function z is going to be 1 over the square root of 1 minus x squared. Such things come up pretty often so we have to know its derivative. We could graph it. That's a perfectly reasonable function, and it's a perfect chain. The first point is to identify what's the inside function and what's the outside.

So inside I'm seeing this 1 minus x squared. That's the quantity that it'll be much simpler if I just give that a single name y. And then what's the outside function? What am I doing to this y? I'm taking its square root, so I have y to the 1/2. But that square root is in the denominator. I'm dividing, so it's y to the minus 1/2. So z is y to the minus 1/2.

OK, those are functions I'm totally happy with. The derivative is what? dz dy, I won't repeat the chain rule. You've got that clearly in mind. It's right above. Let's just put in the answer here. dz dy, the derivative, that's y to some power, so I get minus 1/2 times y to what power? I always go one power lower. Here the power is minus 1/2. If I go down by one, I'll have minus 3/2. And then I have to have dy dx, which is easy. dy dx, y is 1 minus x squared. The derivative of that is just minus 2x.

And now I have to assemble these, put them together, and get rid of the y. So the minus 2 cancels the minus 1/2. That's nice. I have an x still here, and I have y to the minus 3/2. What's that? I know what y is, 1 minus x squared, and so it's that to the minus 3/2. I could write it that way. x times 1 minus x squared-- that's the y-- to the power minus 3/2. Maybe you like it that way. I'm totally OK with that. Or maybe you want to see it as-- this minus exponent down here as 3/2. Either way, both good.

OK, so that's one more practice. and I've got one more in mind. But let me return to this board, the starting board, just to justify where did this chain rule come from. OK, where do derivatives come from? Derivative always start with small finite steps, with delta rather than d. So I start here, I make a change in x, and I want to know the change in z. These are small, but not zero, not darn small.

OK, all right, those are true quantities, and for those, I'm perfectly entitled to divide and multiply by the change in y because there will be a change in y. When I change x, that produces a change in g of x. You remember this was the y. So this factor-- well, first of all, that's simply a true statement for fractions. But it's the right way. It's the way we want it. Because now when I show it, and in words, it says when I change x a little, that produces a change in y, and the change in y produces a change in z. And it's the ratio that we're after, the ratio between the original change and the final change. So I just put the inside change up and divide and multiply.

OK, what am I going to do? What I always do, whatever body does with derivatives at an instant, at a point. Let delta x go to 0. Now as delta x goes to 0, delta y will go to 0, delta z will go to 0, and we get a lot of zeroes over 0. That's what calculus is prepared to live with. Because it keeps this ratio. It doesn't separately think about 0 and then later 0. It's looking at the ratio as things happen. And that ratio does approach that. That was the definition of the derivative. This ratio approaches that, and we get the answer. This ratio approaches the derivative we're after. That in a nutshell is the thinking behind the chain rule.

OK, I could discuss it further, but that's the essence of it. OK, now I'm ready to do one more example that isn't just so made up. It's an important one. And it's one I haven't tackled before. My function is going to be e to the minus x squared over 2. That's my function. Shall I call it z? That's my function of x. So I want you to identify the inside function and the outside function in that change, take the derivative, and then let's look at the graph for this one. The graph of this one is a familiar important graph. But it's quite an interesting function.

OK, so what are you going to take? This often happens. We have e to the something, e to some function. So that's our inside function up there. Our function y, inside function, is going to be minus x squared over 2, that quantity that's sitting up there. And then z, the outside function, is just e to the y, right? So two very, very simple functions have gone into this chain and produced this e to the minus x squared over 2 function.

OK, I'm going to ask you for the derivative, and you're going to do it. No problem. So dz dx, let's use the chain rule. Again, it's sitting right above. dz dy, so I'm going to take the slope, the derivative of the outside function dz dy, which is e to the y. And that has that remarkable property, which is why we care about it, why we named it, why we created it. The derivative of that is itself.

And the derivative of minus x squared over 2 is-- that's a picnic, right?-- is a minus. x squared, we'll bring down a 2. Cancel that 2, it'll be minus x. That's the derivative of minus x squared over 2. Notice the result is negative. This function is at least out where-- if x is positive, the whole slope is negative, and the graph is going downwards.

And now what's-- everybody knows this final step. I can't leave the answer like that because it's got a y in it. I have to put in what y is, and it is-- can I write the minus x first? Because it's easier to write it in front of this e to the y, which is e to the minus x squared over 2. So that's the derivative we wanted. Now I want to think about that

function a bit.

OK, notice that we started with an e to the minus something, and we ended with an e to the minus something with other factors. This is typical for exponentials. Exponentials, the derivative stays with that exponent. We could even take the derivative of that, and we would again have some expression. Well, let's do it in a minute, the derivative of that.

OK, I'd like to graph these functions, the original function z and the slope of the z function. OK, so let's see. x can have any sign. x can go for this-- now, I'm graphing this. OK, so what do I expect? I can certainly figure out the point x equals 0. At x equals 0, I have e to the 0, which is 1. So at x equals 0, it's 1.

OK, now at x equals to 1, it has dropped to something. And also at x equals minus 1, notice the symmetry. This function is going to be-- this graph is going to be symmetric around the y-axis because I've got x squared. The right official name for that is we have an even function. It's even when it's same for x and for minus x.

OK, so what's happening at x equal 1? That's e to the minus 1/2. Whew! I should have looked ahead to figure out what that number is. Whatever. It's smaller than 1, certainly, because it's e to the minus something. So let me put it there, and it'll be here, too.

And now rather than a particular value, what's your impression of the whole graph? The whole graph is--It's symmetric, so it's going to start like this, and it's going to start sinking. And then it's going to sink. Let me try to get through that point. Look here. As x gets large, say x is even just 3 or 4 or 1000, I'm squaring it, so I'm getting 9 or 16 or 1000000. And then divide by 2. No problem.

And then e to the minus is-- I mean, so e to the thousandth would be off that board by miles. e to the minus 1000 is a very small number and getting smaller fast. So this is going to get-- but never touches 0, so it's going to-- well, let's see. I want to make it symmetric, and then I want to somehow I made it touch because this darn finite chalk. I couldn't leave a little space. But to your eye it touches. If we had even fine print, you couldn't see that distance.

So this is that curve, which was meant to be symmetric, is the famous bell-shaped curve. It's the most important curve for gamblers, for mathematicians who work in probability. That bell-shaped curve will come up, and you'll see in a later lecture a connection between how calculus enters in probability, and it enters for this function.

OK, now what's this slope? What's the slope of that function? Again, symmetric, or maybe anti-symmetric, because I have this factor x. So what's the slope? The slope starts at 0. So here's x again. I'm graphing now the slope, so this was z. Now I'm going to graph the slope of this.

OK, the slope starts out at 0, as we see from this picture. Now we can see, as I go forward here, the slope is

always negative. The slope is going down. Here it starts out-- yeah, so the slope is 0 there. The slope is becoming more and more negative. Let's see. The slope is becoming more and more negative, maybe up to some point. Actually, I believe it's that point where the slope is becoming-- then it becomes less negative. It's always negative. I think that the slope goes down to that point x equals 1, and that's where the slope is as steep as it gets. And then the slope comes up again, but the slope never gets to 0. We're always going downhill, but very slightly.

Oh, well, of course, I expect to be close to that line because this e to the minus x squared over 2 is getting so small. And then over here, I think this will be symmetric. Here the slopes are positive. Ah! Look at that! Here we had an even function, symmetric across 0. Here its slope turns out to be-- and this could not be an accident. Its slope turns out to be an odd function, anti-symmetric across 0. Now, it just was. This is an odd function, because if I change x, I change the sign of that function.

OK, now if you will give me another moment, I'll ask you about the second derivative. Maybe this is the first time we've done the second derivative. What do you think the second derivative is? The second derivative is the derivative of the derivative, the slope of the slope. My classical calculus problem starts with function one, produces function two, height to slope. Now when I take another derivative, I'm starting with this function one, and over here will be a function two.

So this was dz dx, and now here is going to be the second derivative. Second derivative. And we'll give it a nice notation, nice symbol. It's not dz dx, all squared. That's not what I'm doing. I'm taking the derivative of this. So I'm taking-- well, the derivative of that, I could-- I'm going to give a whole return to the second derivative. It's a big deal. I'll just say how I write it: dz dx squared. That's the second derivative. It's the slope of this function.

And I guess what I want is would you know how to take the slope of that function? Can we just think what would go into that, and I'll put it here? Let me put that function here. minus x e to the minus x squared over 2. Slope of that, derivative of that. What do I see there? I see a product. I see that times that. So I'm going to use the product rule. But then I also see that in this factor, in this minus x squared over 2, I see a chain. In fact, it's exactly my original chain. I know how to deal with that chain.

So I'm going to use the product rule and the chain rule. And that's the point that once we have our list of rules, these are now what we might call four simple rules. We know those guys: sum, difference, product, quotient. And now we're doing the chain rule, but we have to be prepared as here for a product, and then one of these factors is a chain.

All right, can we do it? So the derivative, slope. Well, slope of slope, because this was the original slope. OK, so it's the first factor times the derivative of the second factor. And that's the chain, but that's the one we've already

done. So the derivative of that is what we already computed, and what was it? It was that. So the second factor was minus x e to the minus x squared over 2. So this is-- can I just like remember this is f dg dx in the product rule. And the product, this is-- here is a product of f times g. So f times dg dx, and now I need g times-- this was g, and this is df dx, or it will be. What's df dx? Phooey on this old example. Gone.

OK, df dx, well, f is minus x. df dx is just minus 1. Simple. All right, put the pieces together. We have, as I expected we would, everything has this factor e to the minus x squared over 2. That's controlling everything, but the question is what's it-- so here we have a minus 1; is that right? And here, we have a plus x squared. So I think we have x squared minus 1 times that.

OK, so we computed a second derivative. Ha! Two things I want to do, one with this example. The second derivative will switch sign. If I graph the darn thing-- suppose I tried to graph that? When x is 0, this thing is negative. What is that telling me? So this is the second group. This is telling me that the slope is going downwards at the start. I see it.

But then at x equal 1, that second derivative, because of this x squared minus 1 factor, is up to 0. It's going to take time with this second derivative. That's the slope of the slope. That's this point here. Here is the slope. Now, at that point, its slope is 0. And after that point, its slope is upwards. We're getting something like this. The slope of the slope, and it'll go evenly upwards, and then so on.

Ha! You see that we've got the derivative code, the slope, but we've got a little more thinking to do for the slope of the slope, the rate of change of the rate of change. Then you really have calculus straight. And a challenge that I don't want to try right now would be what's the chain rule for the second derivative? Ha! I'll leave that as a challenge for professors who might or might not be able to do it.

OK, we've introduced the second derivative here at the end of a lecture. The key central idea of the lecture was the chain rule to give us that derivative. Good! Thank you.