

**ANNOUNCER:** 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 [mitopencourseware@ocw.mit.edu](mailto:mitopencourseware@ocw.mit.edu).

**PROFESSOR:** Hi. Welcome once again to our lectures in Calculus Revisited, where today we are going to talk about the calculus of composite functions. Now, recall that we have already mentioned in previous lectures the notion of a composite function. And what we're going to do today is to emphasize the idea as to how often we are called upon to find functional relationships, where the first variable is given in terms of a second variable, and the second variable, say, is given in terms of the third variable. And we wish to find, say, the first variable in terms of the third. Fact here is where the name "the chain rule" seems to come from, a chain reaction where the variables are related in a chain this way.

Now, we can see this quite easily in terms of a diagram. Suppose, for example, that I have a graph of 'x' versus 'y'. And suppose also, I have a graph of 't' versus 'x'. Without any reference to calculus-- and this is rather important-- without any reference to calculus, notice that these two graphs together allow me to visualize 'y' in terms of 't'. For example, given a particular value of 't'-- let's call it 't sub 1'-- given a value of 't', from that value of 't', I can find the corresponding value of 'x'. Let's call that 'x sub 1'.

Now, knowing 'x sub 1', I can come to this diagram. Knowing what 'x sub 1' is, I can find 'y sub 1'. And so you see in this chain of two diagrams, a particular value of 't' allows me to find a particular value of 'y'. And in this particular way, I can visualize 'y' as a function of 't'. And you see at this stage of the game, there is absolutely no need to have to have any knowledge of calculus to understand what it is that we're discussing. The place that calculus comes in is in the following way.

Let's suppose it happened in this first diagram that the graph of 'y' versus 'x' was smooth. In other words, let's assume that 'y' is a differentiable function of 'x'. In particular, the way I've drawn this diagram here, we're saying suppose 'dy dx' evaluated at 'x' equals 'x1' happens to exist. And suppose also that this graph of 'x' versus 't', this also happens to be a smooth curve-- in other words, that 'x' is a differentiable function of 't'. Again, in the language of calculus, what we're saying is the slope of this curve exists at this particular point, and it's

given by the  $\frac{dx}{dt}$  evaluated at  $t = t_1$ .

Now, without going into a proof at this stage, all we're saying is this. We suspect that if  $y$  is a differentiable function of  $x$ , and  $x$  is a differentiable function of  $t$ , that therefore,  $y$  should also be a differentiable function of  $t$ . Notice it's not a conjecture at all that if  $y$  is a function of  $x$  and  $x$  is a function of  $t$ , that  $y$  is a function of  $t$ . That part is clear. The conjecture is that we suspect that if  $y$  is a differentiable function of  $x$ , and  $x$  is a differentiable function of  $t$ , that  $y$  will be a differentiable function of  $t$ . In still other words, our suspicion is perhaps that a differentiable function of a differentiable function is again a differentiable function.

But even more to the point, not only do we suspect, for example, that the  $\frac{dy}{dt}$  exists here when  $t = t_1$ , but in line with our lecture of last time, we might even begin to suspect, in terms of this fractional notation, that not only does the  $\frac{dy}{dt}$  exist at  $t = t_1$ , but it can be found by multiplying  $\frac{dy}{dx}$  evaluated at  $x = x_1$  by the  $\frac{dx}{dt}$  evaluated at  $t = t_1$ . Again, almost as if the  $dx$  from the numerator here canceled with the  $dx$  from the denominator here, the same as what we hope our differential notation would be. The question is granted that we would like a result like this to hold true, in a course such as calculus, where we're working with very tiny numbers and quotients of small numbers, places where we've seen that our intuition often leads us astray, it becomes fairly apparent that we had better have something stronger than just intuition in helping us derive certain results, no matter how natural these results might look.

Now, the way we proceed here is as follows again, and notice again the building blocks of calculus. We go back to the fundamental result of last time. You see, after all, to find  $\frac{dy}{dt}$ , we want  $\frac{\Delta y}{\Delta t}$ , and then we'll take the limit as  $\Delta t$  approaches 0. The question is, first of all, do we have a nice expression for  $\Delta y$ ? And in terms of the lecture of last time, we saw that if  $y$  was a differentiable function of  $x$ , that  $x = x_1$ , that  $\Delta y$  was given by  $\frac{dy}{dx}$ , evaluated  $x = x_1$  times  $\Delta x$  plus  $k$  times  $\Delta x$ -- and this is crucial now-- where the limit of  $k$  as  $\Delta x$  approaches 0 was 0.

Now you see, this recipe here is ironclad. I emphasized it from a geometric point of view last time, but you may recall that I proved it from an analytical point of view. In other words, whether you want to visualize this or derive it, it makes no difference. The key factors that this statement here is ironclad. It's something that we now know to be true in our so-called game of calculus.

The point is, again, now how do we use this to check over our conjectured result? Again, the answer is almost straightforward. If you keep track of these things, you'll notice that calculus is a one-step-at-a-time procedure. Namely, we want  $\frac{dy}{dt}$ . That suggests we first want  $\frac{\Delta y}{\Delta t}$ , and then we'll take the limit as  $\Delta t$  approaches 0. So first, we do this. Namely, starting with our known recipe, we divide through by  $\Delta t$ , and why can we do this? We can do this because, of course,  $\Delta t$  is not 0.

Now we take the limit of both sides of the equality as  $\Delta t$  approaches 0. We observe that on the left-hand side, the limit of  $\frac{\Delta y}{\Delta t}$ , as  $\Delta t$  approaches 0, is precisely  $\frac{dy}{dt}$ , and in this particular case, evaluated at  $t$  equals  $t_1$ . In other words, notice that the left-hand side here, as we let  $\Delta t$  approach 0, becomes the left-hand side of our conjecture. Now we recall again that the limit of the sum is the sum of the limits, and we now take the limit of each of these terms separately, each term as a product. The limit of a product is the product of the limits.  $\frac{dy}{dx}$  evaluated at  $x$  equals  $x_1$  is a constant. In fact, that's just what, it's  $\frac{dy}{dx}$ . The limit of  $\frac{dy}{dx}$  evaluated  $x$  equals  $x_1$ , as  $\Delta t$  approaches 0, is just  $\frac{dy}{dx}$  evaluated  $x$  equals  $x_1$ .

On the other hand, by definition, the limit of  $\frac{\Delta x}{\Delta t}$ , as  $\Delta t$  approaches 0, is just  $\frac{dx}{dt}$ . And keeping track of the subscripts here, later on we'll become sloppy and leave the subscripts out. There really is no great harm done in calculus of a single variable. We shall find, in calculus of several variables, that it is extremely important to keep track of the subscripts and where the variables are being evaluated and things of this particular type. But I just want to get you used to the fact that these are specific numbers that we're using over here. Now let's continue.

We take the limit of this term as  $\Delta t$  approaches 0. We observe that this becomes  $\frac{dx}{dt}$ , and the limit of  $k$  as  $\Delta t$  approaches 0-- well, as  $\Delta t$  approaches 0, the fact that  $x$  is a differentiable function of  $t$  means that  $\Delta x$  approaches 0. And since the limit of  $k$  as  $\Delta x$  approaches 0 is 0, this term becomes 0. 0 times anything is, any finite number, is 0. That means that this term here in the limit becomes 0, and we're left with the desired result.

But notice that we did not arrive at this desired result by hand waving. We did not say this term  $\frac{\Delta x}{\Delta t}$  is getting small, so it's becoming negligible. I can't emphasize this point enough that it is true that  $\Delta x$  is becoming small here, but so is  $\Delta t$ , and that indicates, essentially, you're 0 over 0 form. And the thing that saves us, the thing that makes this whole term drop out, is the key fact that  $k$  itself goes to 0, as  $\Delta x$  goes to 0.

By the way, there are easier ways of intuitively trying to remember the chain rule. For example, one way that people often try to visualize the chain rule is this. They'll say, OK, we want  $dy/dt$ . So let's take  $\Delta y$  divided by  $\Delta t$ , and then we'll take the limit as  $\Delta t$  approaches 0. Now, you see in this notation here,  $\Delta y$  and  $\Delta t$  are actually numbers. As numbers, we can write these things in fractional notation, and we could write, what, that  $\Delta y$  divided by  $\Delta t$  is  $\Delta y$  divided by  $\Delta x$  times  $\Delta x$  divided by  $\Delta t$ . Then we could take the limit, as  $\Delta t$  approaches 0, and we would arrive at the same result.

But again, without trying to make this thing too obnoxiously long here, the thing to keep in mind is that  $x$  is a function of  $t$ . And from a rigorous point of view, the danger with this shortcut technique-- and it can be patched up but requires a great deal of mathematical analysis-- the danger here is that as  $\Delta t$  approaches 0, it's quite possible that  $\Delta x$  will be 0. In other words, it's possible that for a given change in  $t$ , there is no change in  $x$ . Now, if  $\Delta x$  happens to equal 0, then we're in trouble over here. In other words, in many cases, this shortened version gives us an idea as to what's going on. But our so-called longer method has no pitfalls to it.

But enough said for what this recipe is. This result is known as the chain rule, and this will be the topic of the rest of today's lecture. Now, let's take a look at some of these things in a bit more detail. For example, let's look at an illustration. Suppose we want to find  $dy/dx$ , if  $y$  is equal to  $x^2 + 1$ .

Let me first do this problem the wrong way. Let's put a question mark over here. People learn things like, what, bring the exponent down and replace it by one less. Now certainly, if I bring the exponent down here, and replace it by one less, this is the answer that I get. Of course the question is, is this the right answer? Well, you see, notice one very nice way about finding out whether an answer is wrong, is to first find out by another way which is the right answer.

For example, if  $y$  equals  $x^2 + 1$ , it happens that we know how to square this thing. We can find directly that another way of expressing  $y$  is what? It's  $x^4$  plus  $2x^2$  plus 1, but we have previously learned how to differentiate a polynomial. Through the polynomial is what, this is going to be what?  $4x^3$  plus  $4x$ .

And you see somehow or other, this does not seem to give-- well, for one thing, we see that these are two different answers. For another thing, if this is the one that happens to be the right answer, this is the one that is the wrong answer. And since we know from previous

material this is the right answer, there is something wrong with this regardless of how right it might look. In fact, how much are we off by over here? If we factor this thing out, what can we do? We can write this as  $4x$  times  $x^2 + 1$ . And what we really had over here was twice  $x^2 + 1$ . It seems that the correction factor is  $2x$ . Now again, notice that the derivative of what's inside the parentheses over here just happens to be exactly  $2x$ .

Now how does the chain rule come into play in a problem of this type? You see, the thing is, that what we should do over here is rewrite this. Namely, for example, let  $u$  equal  $x^2 + 1$ . Then what this says is what?  $y$  is equal to  $u^2$ , where  $u$  is equal to  $x^2 + 1$ . This is just another way of writing this, and in this particular form, the chain rule seems to be emphasized more. You see,  $y$  is a function of  $u$ ,  $u$  is a function of  $x$ .

Notice that from the first equation, it is relatively easy to find  $dy/du$ . In fact, it's just what,  $2u$ . We'll write that down later. From the second equation, it's easy to find  $du/dx$ . And by the chain rule, all we're saying is that  $dy/du$  times  $du/dx$  is  $dy/dx$ . See, what will that give us in this case?  $dy/du$  is  $2u$ .  $du/dx$  is  $2x$ . That gives us  $4x$  times  $u$ .  $u$  is  $x^2 + 1$ , and so this becomes  $4x$  times  $x^2 + 1$ . And if we now compare this with what was the correct answer, we see that in this case, everything worked out fine.

I suppose what we should do here is to comment now on the danger of memorizing recipes without thoroughly understanding them. The idea, that said when you want to differentiate something raised to a power that you bring the power down and replace it by one less, hinged on the fact that the thing that was being raised to the power was the same variable with respect to which you were doing the differentiation.

You see, for example, when we had  $y$  equal to  $x^2$ , and then we wrote that  $dy/dx$  is  $2x$ , the thing that was important over here was the fact that what? The thing that was being raised to the second power is precisely the variable with respect to which we were doing the differentiation. You see, in the problem  $y$  equals  $(x^2 + 1)^2$ , the thing that was being raised to the second power was  $x^2 + 1$ . The variable with respect to which we were differentiating was  $x$ .

In other words, to write this thing more symbolically, if  $y$  is equal to something, square it, then the derivative that's equal to twice that something is the derivative of  $y$  with respect to that something. You see, the place the chain rule comes in is when the variable which appears here, is not the same as the variable which appears here, and we'll see this in greater detail as

we go along.

By the way, the chain rule comes up in another form known as parametric equations, and this form comes up very often. It's a twist of what we were talking about before. This is the situation in which frequently we want to compare two variables. Let's call them 'x' and 'y', all right? And it happens that both variables, 'x' and 'y', can be expressed more simply in terms of a third variable, 't'. And frequently, what one does is try to talk about the relationship that exists between 'y' and 'x' in terms of eliminating t between these two equations.

By the way, in terms of differential language, there seems to be an easier way of handling this. Namely, you see, if we differentiate the first equation, we get what, that 'dy dt' is 'f prime of t'. If we differentiate the second equation, we get that 'dx dt' is 'g prime of t'. Now if, as we said in our last lecture, we can pretend that this is really a fraction, that it's 'dy' divided by 'dt'-- in other words, if we think of 'dy' as being 'delta y-tan', of 'dx' as being 'delta x-tan', and 'dt' as being 'delta t', it would appear that we could say, what, that 'dy dt' divided by 'dx dt' would just be what? 'dy dx'. In other words, "dy' divided by 'dt'" divided by "dx' divided by 'dt'", which is what this would say if this was in differential form, would just be 'dy dx'. In other words, we get the feeling that to find the derivative here, all we have to do is differentiate 'y' with respect to 't', and divide that by the derivative of 'x' with respect to 't'.

And by the way, you see, this becomes a particularly powerful tool in those computational cases where we do not know how to eliminate 't', and to solve specifically for 'y' in terms of 'x'. You see, in terms of this particular recipe over here, we are allowed to leave 'x' and 'y' in terms of 't'. Again, the same old bugaboo comes up to plague us. The fact that something seems natural is not enough to allow us to believe that it's actually correct.

Is there a more rigorous way of obtaining the same result? Again, the answer is yes. And not only is the answer yes, but it goes back to the fundamental recipe that we were discussing in our previous lecture. Namely, we know that 'delta y' is "f prime of t' times 'delta t'" plus 'k1 delta t', and the 'delta x' is "g prime of t' times 'delta t'" plus 'k2 delta t', where both the limit of 'k1' and 'k2' as 'delta t' approach 0. And this is a notation, I think, that takes a while to get used to. We're used to seeing letters like 'k' stand for constants, but it's important over here to understand that 'k1' and 'k2' are functions of 'delta t', that the difference between 'delta y' and 'delta y-tan', 'delta x' and 'delta x-tan', that difference, which is 'k delta x' or 'k delta y', depending on which problem we're dealing with that's certainly 'k' in that case, does depend on how big 'delta t' happens to be.

At any rate, the important thing is that as ' $\Delta t$ ' approaches 0, these go to 0 also. Now you see if we take this, and actually compute ' $\Delta y$ ' divided by ' $\Delta x$ '-- and we'll write this a little bit more suggestively, factor out ' $\Delta t$ ' from both numerator and denominator-- it rigorously tells us what ' $\Delta y$ ' divided by ' $\Delta x$ ' is. Now we take the limit, as ' $\Delta x$ ' approaches 0. That, by definition, is what? That's by definition ' $dy dx$ '.

Well, you see, first of all, we cancel out the ' $\Delta t$ ' over here, see, ' $\Delta t$ ' is not 0, we're assuming. Since it's not 0 it can be canceled out, and once we've canceled out ' $\Delta t$ ', notice that as ' $\Delta t$ ' approaches 0, so does ' $\Delta x$ '. As ' $\Delta x$ ' approaches 0, so does ' $\Delta t$ '. That makes ' $k_1$ ' and ' $k_2$ ' go to 0. And then since the limit of a quotient is the quotient of the limits, provided only the ' $g$  prime of  $t$ ' is not 0, we see that in the eliminating process, we get the same answer.

And by the way, see, once we get the same answer, as we would have got the short way, then we can use the convenience of the short recipe. However, the fact that the short recipe was nice is not enough of a guarantee that it was giving us the correct answer. As a case in point, it's rather interesting to point out that if you want the second derivative-- in other words, let's recall what we have here. We have what? ' $y$ ' was given to ' $b$ ', say ' $f$  of  $t$ '. ' $x$ ' was given by ' $g$  of  $t$ '. And you see from these two equations, what we could do is find what? We could find the second derivative of ' $y$ ' with respect to  $t$ ', and we could also find from this equation the second derivative of ' $x$ ' with respect to ' $t$ '.

This we could certainly do. And mechanically, we could certainly say, let's cancel the common denominator. The interesting thing is that when you form that quotient, whatever that quotient is, it does not come out to be the second derivative of ' $y$ ' with respect to ' $x$ '.

And there is an interesting piece of folklore over here. I don't know if this ever bothered you or not, but it used to bother me. I never understood why, when you talk about the second derivative, that the exponent was written between the ' $d$ ' and the variable in one case, but written at the end in the other case. In other words, notice that the 2 here appears between the ' $d$ ' and the ' $y$ ', but in the denominator, the ' $d$ ' appears outside. And again, it was the foresight of the fathers of differential calculus who noticed rather interestingly that if mechanically you did agree to cancel the common denominator here, that what you would wind up with is not ' $d^2y dx$  squared', but rather what? ' $d^2y d^2x$ '. In other words, if you mechanically carried this out, notice that the notation would be incorrect. The 2 comes out to

be in the wrong place over here.

You see, again, the interesting point is we don't have to rely on taking my word for it.

Somebody might say to me, now look, all you've told me is that I get the wrong answer solving this problem this particular way. And you've given me a nice lecture about how the 2's come out the wrong way and everything. How do I know that this is the wrong answer?

See, and again, everything comes back to fundamentals again. To find  $d^2y/dx^2$ , observe that by definition, that's just  $d/dx$  of  $dy/dx$ . That definition doesn't depend on what functions we're dealing with. The second derivative with respect to  $x$  is the derivative with respect to  $x$  of the first derivative. Now, once we have this, you see, knowing from our previous case, that what?  $dy/dx$  was  $f'(t)/g'(t)$ . We can now do what? Take this derivative.

By the way, again, notice how the chain rule comes up in practice. It's not always dictated to us. If you look at the expression inside the parentheses, what do we have? Inside the parentheses, we have a function of  $t$  only. This is a function of  $t$ . We want to differentiate it with respect to  $x$ . The most natural variable to differentiate a function of  $t$  with respect to is  $t$  itself.

In other words, what would've been nice is if this was the derivative of  $f'(t)$  over  $g'(t)$ , with respect to  $t$ . See, this would be easier to handle. We would then use the quotient rule, et cetera. You see, we can differentiate a function of  $t$  with respect to  $t$ . The trouble is we have the derivative with respect to  $x$ . And if we just change this to a  $t$ , that's cheating. See, I mean, you pretend you copy it wrong, because it's an easier problem to solve that way.

The beauty of the chain rule is that it allows us to do the problem the easier way, and to doctor up the resulting incorrect answer by the right answer. Namely, you see what we wanted to wind up with here is what, the derivative not with respect to  $t$ , but with respect to  $x$ . And so, by using the chain rule, you see we do what? We take the derivative with respect to  $t$ , multiply that by  $dt/dx$ -- again, mechanically, almost as if these canceled.

But this is the way the chain rule works, and now, you see, I can work this out by the regular quotient rule, which says what? It's the denominator times the derivative of the numerator. See, and I am differentiating out respect to  $t$ , the natural variable, minus the numerator times the derivative of the denominator over the square of the denominator.

Now, that's a mess by itself, meaning, what, computationally, it's not that obvious. I mean, it's quite a bit of work to do here, and then that whole thing must be multiplied by  $dt dx$ . And this, you see, is how one goes around finding the second derivative of  $y$  with respect to  $x$  in terms of parametric equations. And more than once, if you're not careful, you're going to find yourself making serious mistakes, by forgetting to put in this factor of  $dt dx$ .

By the way, an interesting point is that we have not computed  $dt dx$ . We have computed  $dx dt$ . Let's go back here. See,  $x$  was  $g$  of  $t$ . So from that,  $dx dt$  is  $g'$  of  $t$ . And the question is if  $dx dt$  is  $g'$  of  $t$ , how does one find  $dt dx$ ? And again, I think your intuition is going to tell you to just take reciprocals. And again, the question is it's true that this suggests taking reciprocals, but how do we know that we can do this, and if we can do this, what does it really mean? You see, what this is leading into is what's going to be the subject of our lecture next time, called 'Inverse Functions'. And just to give you a preview of what that lecture is about, and how we work things like this, let's take a look at what we mean by inverse functions. Well, we won't even mention it in much detail. But let's take a look and see what's going on over here.

Let's suppose that the first-- and by the way, I've started to abandon using the  $t$  over here all the time. I think those of us in engineering work primarily keep thinking of  $t$  as being time, and you may get the mistaken notion that if the variable isn't time, the thing doesn't work this way. In most cases, physically, the variable that we're interested in will be time. But just for the idea of getting you used to the fact that it makes no difference what the name of the variable is, I've taken the liberty of writing this slightly differently. Namely, I now assume that  $y$  is a differentiable function of  $u$ , and that  $u$  is a differentiable function of  $x$ . By the chain rule, I now know that  $y$  is a differentiable function of  $x$ , and that  $dy dx$  is  $dy du$  times  $du dx$ .

The interesting thing here is, is that there is nothing in the statement of the chain rule that says that the first variable in the third that  $x$  and  $y$  must be different variables. In fact, it might happen that  $x$  and  $y$  are synonyms for one another. If  $x$  and  $y$  happen to be synonyms-- suppose  $x$  and  $y$  are synonyms-- look what happens over here.  $dy dx$  is then just  $dy dy$ , which is 1. See, let's write that down. That's  $dy dy$ . This would be  $dy du$ , and if  $x$  is equal to  $y$ , this is  $du dy$ . And if this is equal to 1, and this is  $dy du$ , and this is  $du dy$ , what does this tell us about the relationship between  $dy du$  and  $du dy$ ? It says their product is 1. And if the product is 1, that by definition means that the two factors are reciprocals.

Now, what I want you to observe over here is what this whole thing means. Namely, if 'y' happens to equal 'x', do you see what this thing says? It says that 'y' is a differentiable function of 'u', and 'u' in turn is a differentiable function of 'y'. That's precisely what we meant when we talked about inverse functions. We don't know when an inverse function exists. All we're saying is, is that if 'f inverse' happens to exist over here, to find 'du dy', all we have to do is take the reciprocal of 'dy du'.

Now again, this is going to be the subject of our next lecture. All I wanted to do was to make this aside for the time being. What I want to do to complete today's lecture is to get to something more tangible. See, now that we've talked about the chain rule, we've talked about inverse functions a little bit, and talked about these things from a highly theoretical point of view, let's go ahead and try to solve a particularly simple problem. By particularly simple, I mean this. I have chosen the numbers to come out in a very, very easy way, so we don't get lost in the maze of details. In other words, there was a danger that we will confuse the computational details with the theory. So to emphasize the theory, I've tried to pick a straightforward simple problem, but let's see how this thing works out.

Let's suppose that we're given that 'y' is equal to 't to the fourth power', and 'x' is equal to 't squared'. What we would like to do-- and by the way, notice what this thing says, a given value of 't' determines both an 'x' and the 'y', so that makes 'x' and 'y' functionally related. Notice that from the first equation, we can find that 'dy dt' is '4 t cubed'. From the second equation, we can find that 'dx dt' is '2t'. And if we now use the chain rule, 'dy dx' will be what? It'll be 'dy dt' divided by 'dx dt', and that's just '2 t squared'.

By the way, as a check, notice this. If 'y' is equal to 't to the fourth', and 'x' is equal to 't squared', since 't to the fourth' is the square of 't squared', that says 'y' is equal to 't squared squared', 'y' is equal to 'x squared'. And if 'y' is equal to 'x squared', in this case, it's very easy to see that 'dy dx' is equal to '2x'. By the way, when we try to compare these two answers, they look different, but that's because they're expressed in terms of different variables. If we return to our original equations, and we see that 'x' is equal to 't squared'-- 'x' is a synonym for 't squared'-- this is the check that we have received the right answer.

By the way, before I conclude today's lecture, I would like to make a rather important aside about parametric equations. After one works the problem this way, and comes down to the check, and says, hey, after all of this mess over here, I could have replaced it by just 'y equals x squared', why did I have to work with this in the first place? We are going to have many,

many examples throughout the course that will illustrate this. But at least once in a lecture, I would like to go on record as pointing out that this pair of equations tells you much more than this equation here.

This equation simply tells you this. If a particle were moving along a curve with respect to time according to these equations, this equation here simply tells you what path the particle would follow. Namely, the parabola 'y' equals 'x squared'. On the other hand, these two equations tell you much more than that. These not only tell you that the particle moved along the parabola 'y' equals 'x squared', but rather, it tells you at a particular time the point on the parabola that the particle was at.

What I mean is this. As another example, suppose we had 'y' equals 't squared,' and 'x' equals 't'. If we eliminate 't' from these two equations, we also find that 'y' is equal to 'x squared'. Yet notice that this is not the same as our original set of equations. For example, here, when 't' is 2, when 't' is 2 over here, what point are we on as far as the parabola is concerned? When 't' is 2, this is 2, and this is 4. That would be the point 2 comma 4.

On the other hand, with respect to this equation, when 't' is 2, 'x' is 4, and 'y' is 16, you see both of these particles would follow the same curve, but they are at different points at different times. So don't belittle the parametric approach. Having the parameter 't' in there tells you more than just what the path of the motion is. It tells you at what time a particle was at what particular point.

Well, enough about that. Let's go ahead and find the second derivative now. You see, we already know that 'dy dx' is '2 t squared'. Now what we'd like to find is 'd<sup>2</sup>y dx squared'. Again, the same basic definition. 'd<sup>2</sup>y dx squared'. The second derivative is the derivative of the first derivative. The first derivative we saw was '2 t squared', so this is the derivative of '2 t squared'.

Again, and this is where most of the mistakes are made, people get sloppy. They forget the 'x' is in here. They say I know the derivative of this, it's '4t'. Well, the derivative of this is '4t' with respect to 't'. We want to differentiate with respect to 'x'. And the way the chain rule comes in, we say OK, since 't' is the natural variable with respect to which to differentiate, let's do it. We'll differentiate in respect to 't'. But since the final answer has to be with respect to 'x', our correction factor by the chain rule will be 'dt dx'.

Well, the derivative of '2 t squared' with respect to 't' is clearly '4t'. The derivative of 't' with

respect to 'x', assuming that we know something about inverse functions, that's the reciprocal of 'dx dt'. We just saw that 'dx dt' was '2t'. Therefore, 'dt dx' is 1 over '2t', and therefore, the correct answer appears to be 2.

Again, this is why I picked the simple case. Given that 'y' equals 'x squared', we see at a glance that 'dy dx' is equal to '2x', and also at a glance, therefore, that 'd<sup>2</sup>y dx squared' is equal to 2. By the way, that's exactly what this is equal to. You see, had we forgot the chain rule, and had we left this factor out, this would have given us-- in other words, to simply write down that the answer was '4t', which is the most common mistake that's made, would have given us the wrong answer. That's why I put such an easy problem. You see, if I had picked a tougher computational problem, the theory would have remained the same. But when I got two different answers, it would have been difficult to determine which was the correct answer, and which was the incorrect answer.

But again, to summarize today's lecture, it was a continuation in a way of the lecture of last time, when we developed the primary recipe involving differentials. Now we applied that to find something called the chain rule. In the process of emphasizing the chain rule, we talked about the necessity of knowing something about inverse functions. Consequently, that dictates what our next lecture will be concerned with, namely inverse functions. And so until next time, goodbye.

**ANNOUNCER:** Funding for the publication of this video was provided by the Gabriella and Paul Rosenbaum Foundation. Help OCW continue to provide free and open access to MIT courses by making a donation at [ocw.mit.edu/donate](https://ocw.mit.edu/donate).