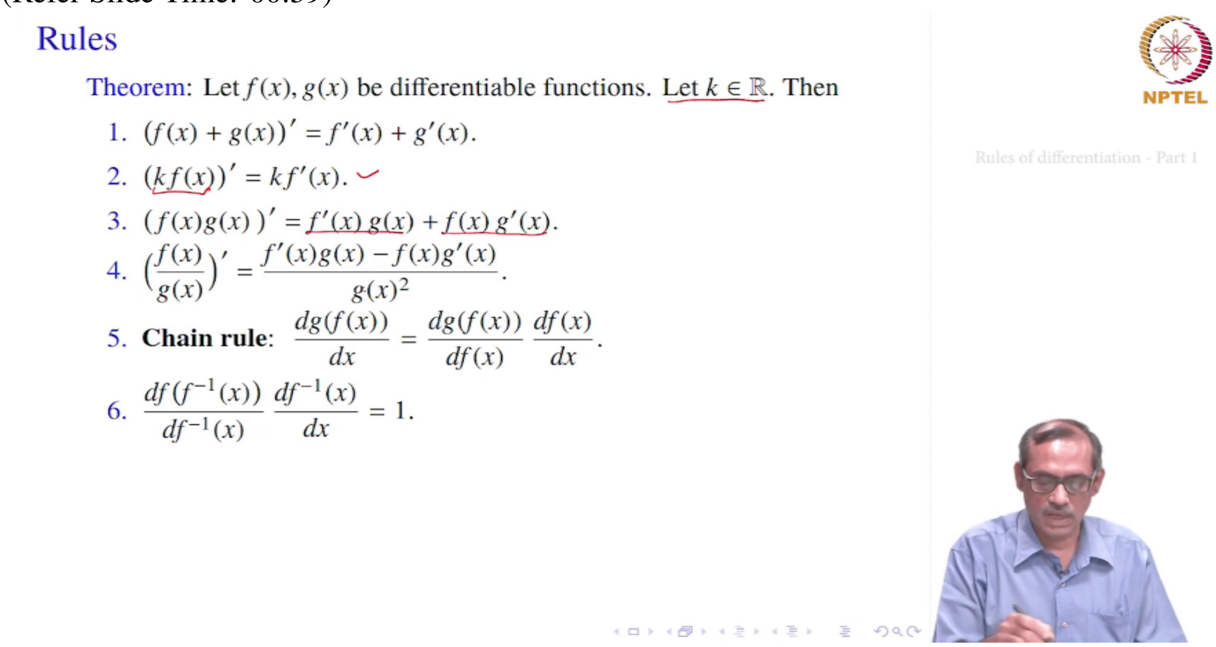
Basic Calculus - 1
Professor. Arindama Singh
Department of Mathematics
Indian Institute of Technology Madras
Lecture 15 - Part 1
Rules of Differentiation - Part 1

This is lecture 15 of Basic Calculus 1. In the last two or three lectures, we had discussed about differentiation, that is, how to take the derivative of a function. Sometimes the function is differentiable at a point, and sometimes it is not. If it is differentiable, then how to compute its derivative? That is what we will be discussing today.

If you remember, there was a problem where $x\sin(1/x)$ was there. We told we will be talking about rules of differentiation so that you can differentiate easily, instead of going back to the first principle. We will be discussing those things today so that it will be easier to differentiate functions.
(Refer Slide Time: 00:59)



So, this is really compelling, and we will be requiring the usual things. Suppose $f(x)$ and $g(x)$ are differentiable functions. That means at every point of their domains, they are differentiable. Moreover, we are going to consider functions like $f(x)+g(x)$. This new function, which is $f+g$ is sometimes written as $f(x)+g(x)$. That should be definable on the same domain.

As a prerequisite, whenever these kinds of things come where both $f(x)$ and $g(x)$ are involved, we would assume implicitly that they have the same domain. Suppose on their domain, same domain, $f(x)$ and $g(x)$ are both differentiable functions. We start with a constant sayk, which is a real number; it is a constant. That will be, of course, required here.

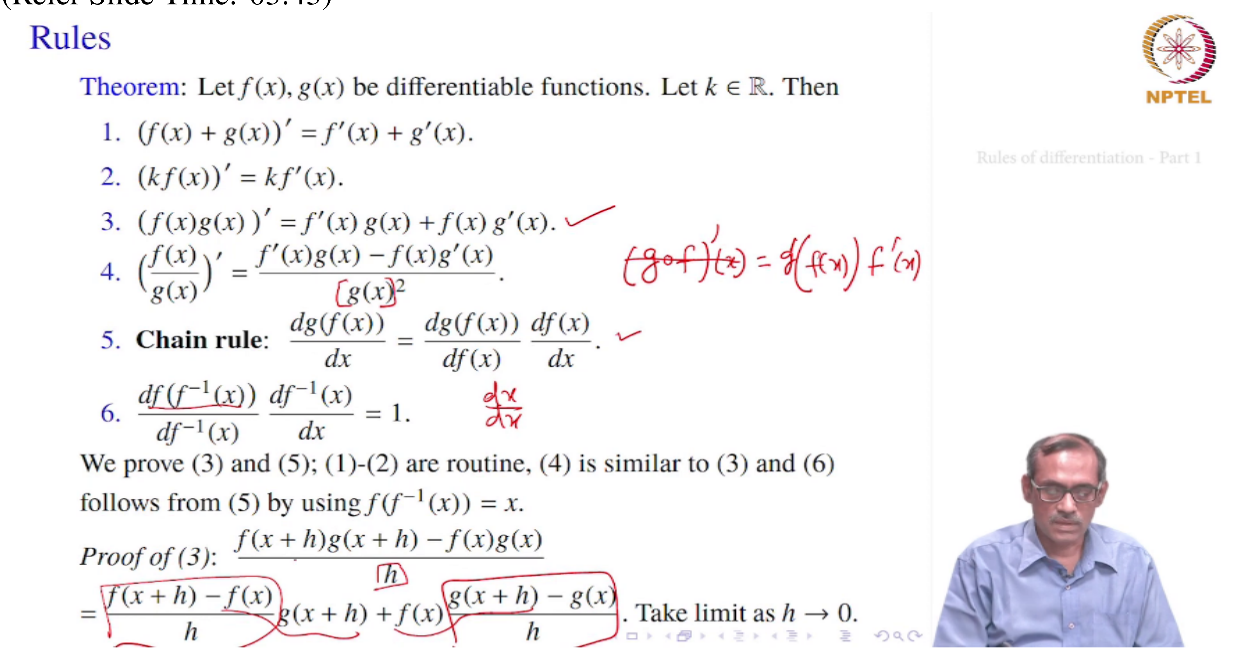
Now, the first rule says that if you take the addition of two functions, that is, $f+g$, (its value at x is of course $f(x)+g(x)$), which we write as the new function $f(x)+g(x)$, then its derivative is equal to sum of the derivatives of individual functions. That is not difficult to see, because you know the algebra of limits. Once you take this increment in $f(x) + g(x)$, it will be in the form $f(x+h)f(x)$. For $g(x)$, it will be $g(x+h) - g(x)$, and for $f(x) + g(x)$ it is $f(x+h) + g(x+h) - f(x) - g(x)$. When you divide it by h and take the limit as $h \to 0$, because of algebra of limits, it will be equal to $f'(x) + g'(x)$.

Similarly, when you take the constant multiple of a function, then in the limiting process that k comes out, and you get k times $f'(x)$.

These two things help us along with the formulas for differentiation of x^n which is equal to nx^{n-1} to differentiate polynomials; we will see of course soon.

The third one says that if you multiply two functions, $f(x)$ and $g(x)$, you get the new function fg, then its derivative is a bit different; it is not just $f'(x)g'(x)$. It will be $f'(x)g(x) + f(x)g'(x)$. You have to remember this and apply. Of course, we will see how it is coming. Well, if you want, let us see it now; this is not difficult.

(Refer Slide Time: 03:43)



We have the derivative of $f(x)g(x)$ on the left side. You have to compute the limit of its increment. The increment in $f(x)g(x)$ is $f(x+h)g(x+h) - f(x)g(x)$. Divide that by the increment in x which is h, and we want to find its limit. Here, for the time being forget this denominator h; we will divide later. Now look at the numerator: $f(x+h)g(x+h) - f(x)g(x)$. We will add and subtract something. Say, we will subtract $f(x)g(x+h)$ and add $f(x+h)g(x+h) - f(x)g(x)$. Now, both the numerators are same. Denominator is h. Once you divide it, you get $[f(x+h) - f(x)]/h$ times $g(x+h)$ plus $[g(x+h) - g(x)]/h$ times $f(x)$. Now, when you take the limit as h goes to 0, the first factor becomes $f'(x)$, and the second one becomes $g'(x)$. So, you get your formula which is in there. So that is how this is proved.

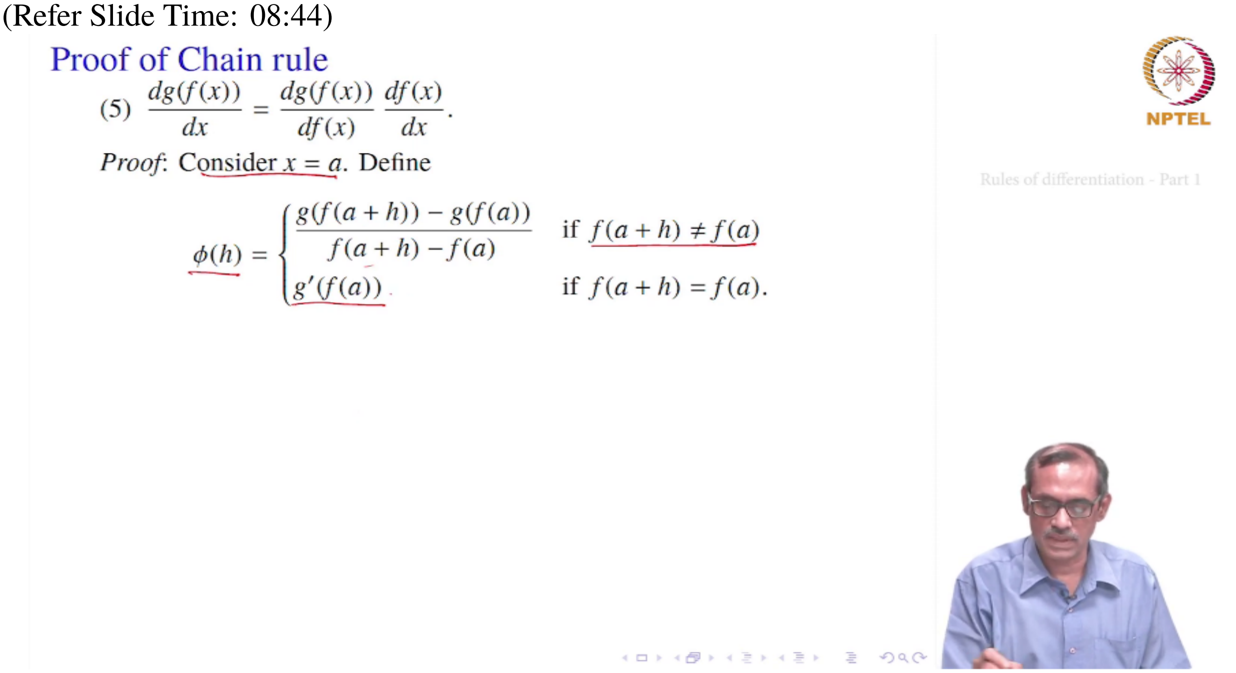
The fourth one follows similarly from the algebra of limits again. That is similar to this; it can be proved. We will go for the general one. But you have to remember this exactly; because once we apply, we do not have to come back here again. It says something about the derivative of

$f(x)/g(x)$. It is not just $f'(x)/g'(x)$ just like the earlier one; it is $f'(x)g(x) - f(x)g'(x)$ divided by $[g(x)]^2$. This is $[g(x)]^2$ not g^2 of x. This is how the formula would look like.

Then the next one is a bit easier to remember, but we will give that later. It says that if you have a composition of functions, $g \circ f$. If you remember, this is defined at x with its value as $g(f(x))$. We are writing on the left side that composition, the derivative of the composition as $(g(f(x))'$; it is equal to the derivative of g with respect to $f(x)$, written as $g'(f(x))$ into $f'(x)$. In $g'(f(x))$, $f(x)$ is a particular number now for x a number so that $f(x)$ is a number. At that point $f(x)$, we take the derivative of g. the symbol syas that it is the derivative of g with respect to f. The answer will be $g'(f(x))$ into $f'(x)$, or $df(x)/dx$. So, this multiplied by $f'(x)$ gives the derivative of the composition: $(g \circ f)' = g'(f)f'$. This is called the chain rule. We will give a proof of this.

In particular, if f is invertible, then it is a bijection so that f^{-1} exists; and you may think of $f \circ f^{-1}$ of x. So, $f(f^{-1}(x)) = x$. On the left side, this first factor is really x, the function is $f \circ f^{-1}$; its value at x is x; so, it is the identity function. That can be differentiated with respect to $f^{-1}(x)$. Then we have $df^{-1}(x)/dx$. This is so because of the chain rule. Look at the right side. It is $df(f^{-1}(x))/dx$; which is equal to dx/dx, and that is equal to 1.

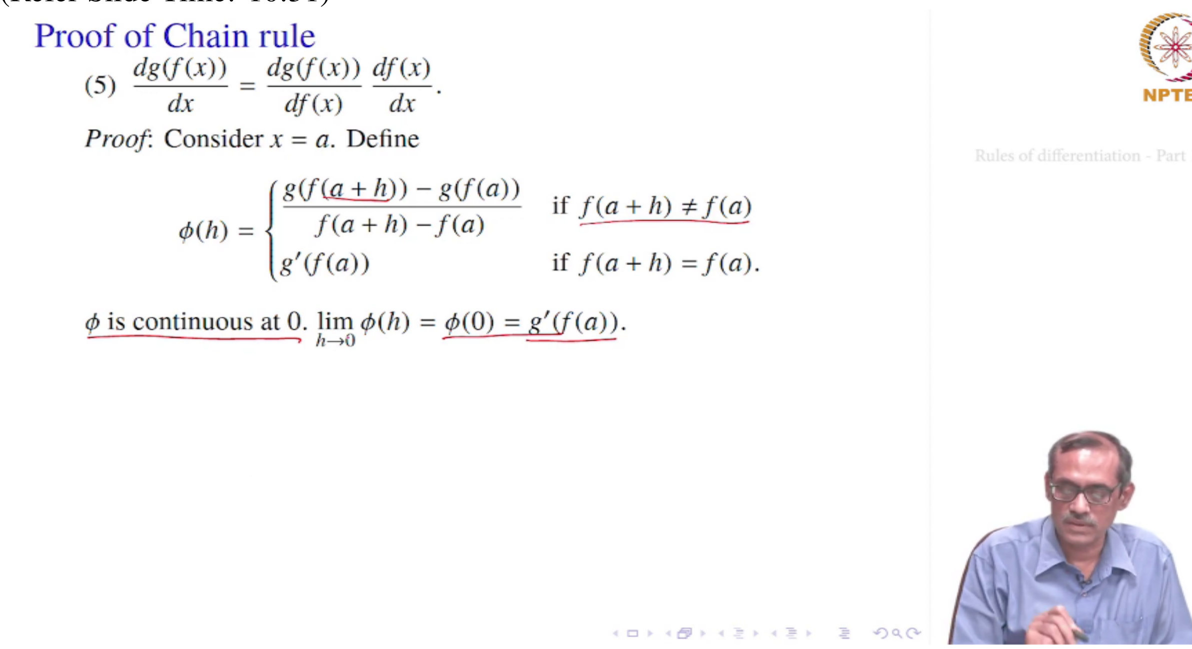
So, sixth one comes out of the fifth by taking in this particular fashion. Or sometimes we write in the other way: $df^{-1}(f(x))/df(x)$ into $f'(x)$. That is also okay. This formula is used to find out the derivative of f^{-1}. We will see it soon how it is applied. But it is coming out of the chain rule. So, let us have a proof of this chain rule. All others are easier; they come from the algebra of limits. (Refer Slide Time: 08:44)



This is the chain rule we want to prove: $d(g(f(x))/dx = dg(f(x))/df(x)$ into $df(x)/dx$. Let x be any point in the domain, sy, $x = a$. We want to differentiate and see that the formula holds at $x = a$; then it will be true for all the points in the domain. So, fix a point, say, $x = a$. At that point a, we want to see that this formula holds. We define a new function $\phi(h) = [g(f(a+h)) - g(f(a))]/[f(a+h) - f(a)]$. Since a is fixed, you can look at this as a function of h. How? You are taking $g(f(a+h))$. This is really a function of h. The increment is $g(f(a+h)) - g(f(a))$. It is a function of h minus a number. So, this is really defining a function of h. We take this when $f(a+h) \neq f(a)$. Of course when it is equal, we cannot define it. So, when it is equal, we define it separately. We say $\phi(h) = g'(f(a))$ when $f(a+h) = f(a)$.

Our assumptions are obvious that g should be differentiable at this point $f(a)$. That is of course there because of the implicit assumption that it is differentiable; then only we can find its derivative. So, we define this function $\phi(h)$, which is equal to this if $f(a+h) \neq f(a)$; and if that is equal, we take $\phi(h)$ as $g'(f(a))$.

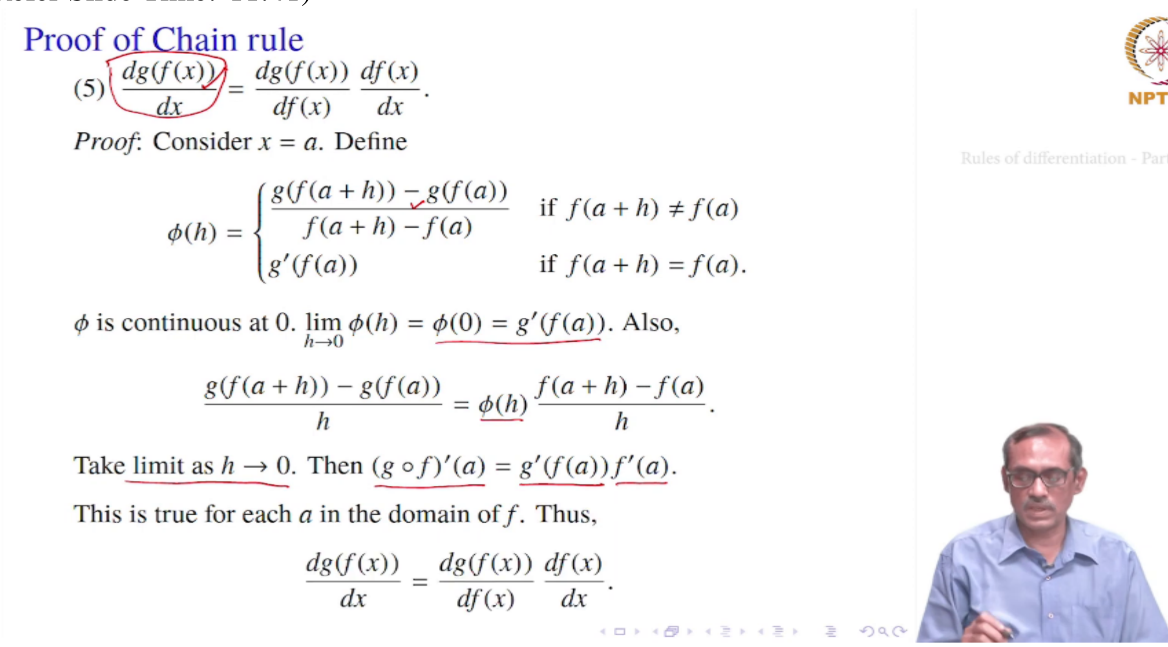
(Refer Slide Time: 10:31)



Now, we want to discover some properties of this ϕ, which depends on h. Our first claim is that ϕ is continuous at 0. Why is it so? It is a function of h. Remember that there is no a now; it is absorbed there as a number. Now, it is continuous at 0 means the limit of $\phi(h)$ as h goes to 0 should be equal to $\phi(0)$. When you take the limit as h goes to 0, that means $h \neq 0$ and $f(a+h) \neq f(a)$. Therefore, this is $\phi(h)$. When you take the limit as h goes to 0, this is really the derivative. It gives the derivative $g'(f(a))$. According to our definition, $\phi(0)$ is equal to $g'(f(a))$. So, the limit of $\phi(h)$ as h goes to 0 is $\phi(0)$. We thus say that ϕ is continuous at 0. This is one property of ϕ; it is continuous at $h = 0$.

Again, we observe that $[g(f(a+h)) - g(f(a))]/h$ in the limit would give us the left side. So that ratio is equal to $\phi(h)$ times $[f(a+h) - f(a)]/h$. How is it true? If $h \neq 0$, then $f(a+h) \neq f(a)$. Because $f(a+h) \neq f(a)$, that applies. Once that applies, this divided by h would give $\phi(h)$ times $[f(a+h) - f(a)]/h$. It is just applying this formula, whenever $f(a+h) \neq f(a)$. If $f(a+h) = f(a)$, then the right side becomes 0. And the left side is, again 0, because $h = 0$. So, this formula is true, in any case.

(Refer Slide Time: 11:41)



We are of course, interested when $h \neq 0$. In that case, you take limit as h goes to 0. The left side gives this limit, which is $g'(f))$ or composition derivative at a: $dg(f(x))/dx$ at $x = a$. And the right side gives $\phi(h)$. The function $\phi(h)$ is continuous. So, that gives you $\phi(0)$, which we know to be $g'(f(a))$. The limit of the second factor as h goes to 0 gives you $f'(a)$. That is how we get the derivative of the composition or the chain rule.

We will see how to apply this chain rule. Of course, its application is for the inverse first. Note that we also write it in this fashion. We have already remarked that.

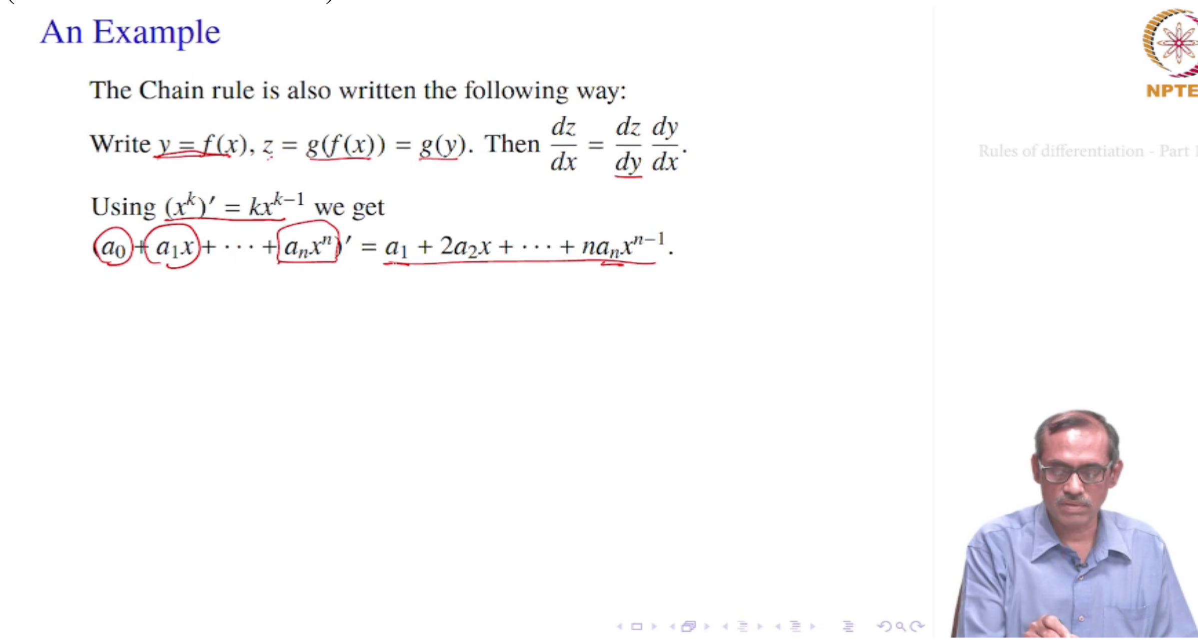
Let us see how this chain rule is applied. First, let us rewrite it. What happens is, we take $y = f(x)$ and write $z = g(f(x)) = g(y)$. Now that z is a function of y and y is a function of x. That is how z becomes a function of x after this composition. The formula now looks like

$$\frac{dg(f(x))}{dx} = \frac{dg(f(x))}{df(x)} \frac{df(x)}{dx}, \quad \text{Or} \quad \frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx}.$$

It is easier to remember this way. As if these two are being canceled; but they are not; remember that they are not canceled. There is nothing called cancellation now, because it is just a notation dz/dy. It is not the ratio of dz with dy. It is one entity, dz/dy; it is not the ratio of dz and dy. But as we have this in the chain rule it looks as if they are ratios. That helps us remembering the formula. But we have to write in this form with $y = f(x)$ and $z = g(y) = g(f(x))$.

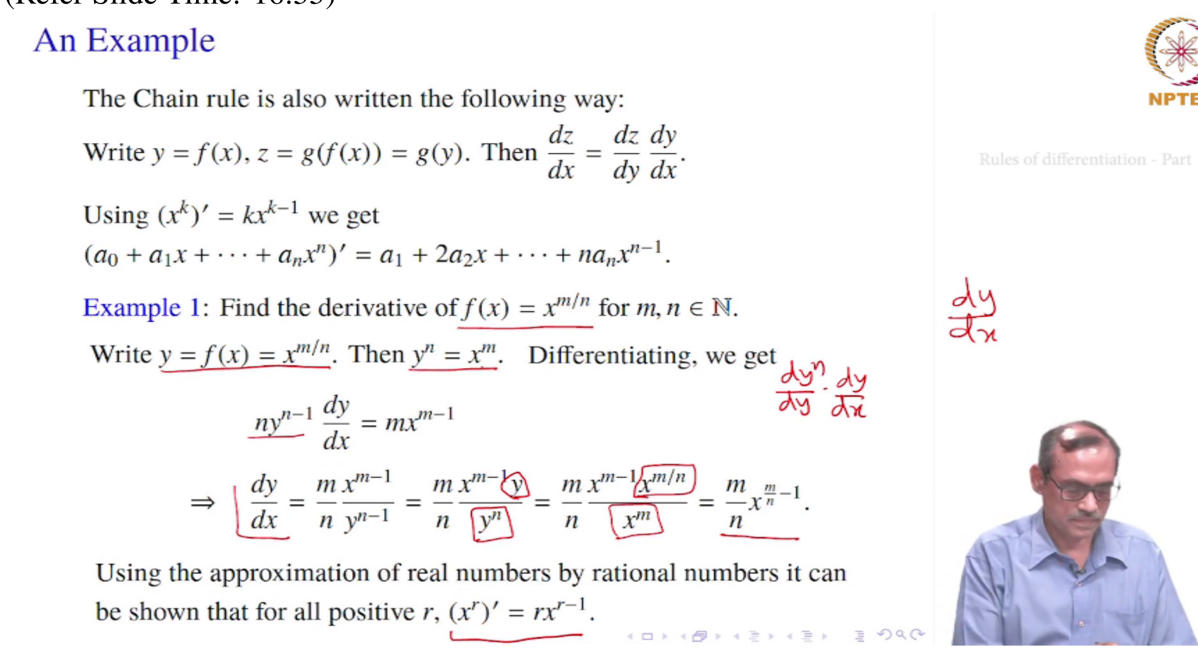
Let us apply these rules to our polynomials. But this is about chain rule. When you differentiate polynomials, you will require only this one, this sum rule, and the constant multiplication rule, the first and second rules. We differentiate a polynomial $a_0 + a_1 x + \cdots + a_n x^n$. It will be the sum of individual derivatives. The derivative of $_0$, which is a constant, is 0. The derivative of $a_1 x$ is a_1 times the derivative of x, which means $k = 1$ here, so, that gives 1 into x^0, which is 1. So, it gives a_1. Similarly, we proceed and the last one is $a_n x^n$. That gives a_n into the derivative of x^n, which is nx^{n-1}. This is how the polynomial is deffierentiated. This is easy to remember, the first constant goes away, next constant remains, next 2 times a_2 and that x^2 comes here and so on.

(Refer Slide Time: 13:43)



Let us see this. We apply chain rule on this function which is $x^{m/n}$. It involves a rational power, not only natural numbers here; we take $x^{m/n}$ where m, n are positive. The negative powers or powers of $1/x$ will come later. So, we have $f(x) = x^{m/n}$. What do we do here? We write $y = f(x) = x^{m/n}$. We want to find out dy/dx. This is what we want to find.

(Refer Slide Time: 16:55)

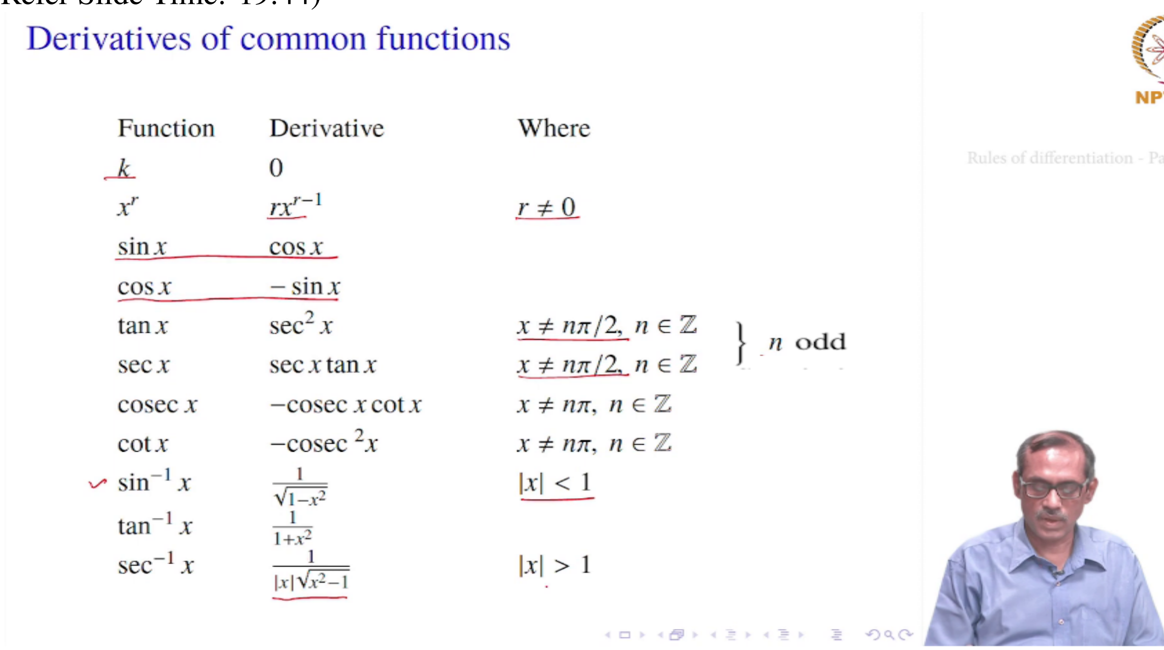


Once $y = x^{m/n}$, we can raise both the sides to power n. It gives $y^n = x^m$. Now, what do we do? We differentiate both the sides so that the derivatives should be same. On the left side you have y^n. Its differentiation will be $d(y^n)/dy$ times dy/dx. That gives the left side as ny^{n-1} times dy/dx. This is on the left side. On the right side, we apply this formula. That gives mx^{m-1}. Now, you have to simplify this substituting our y. So, dy/dx is equal to m/n times x^{m-1}/y^{n-1}. It is a matter of simplification now. $y^{n-1} = y^n$ times y^{-1}, which goes to the top. So, y is here. Now, y^n is replaced by x^m and y is replaced by $x^{m/n}$; all other factors remain. You simply get this: m/n remains as it is, x^m cancels, one x^{-1} remains, and that gives $x^{m/n-1}$.

Look at the final result. It says that $dy/dx = (m/n)x^{m/n-1}$. It looks as if this m/n is something like a k. So, this gives k times x^{k-1}. That means it is the same formula we are using even for k as a rational number. It is easy to remember this way.

Now, this is for rationals. If you take x to the reals, then you have to give some approximation of real numbers by the rationals, which will require the completeness property. And finally, we would be getting the derivative of x^r as rx^{r-1}. for real numbers r, provided x is positive. Because negative to the power irrationals is not defined. It is almost the same formula that holds. It gives $(x^r)' = rx^{r-1}$. That is about the power function.

(Refer Slide Time: 19:44)



If you proceed that way, you would get this list of derivatives. It will help us in computing derivatives of some complicated functions. The first thing is, if it is a constant function, that is, $f(x) = k$ for every x, then its derivative is 0. That is easy; we have done it earlier. Because $f(x+h)$ is k and $f(x)$ is k. So, the numerator becomes 0, and then the limit of that divided by h also becomes 0. So, constant function has derivative as 0.

If it is x^r, then its derivative will be rx^{r-1}, where $r \neq 0$. If $r = 0$, then of course, its derivative is 0. But it looks absurd to apply it here. Of course, it is coming from there. But it is not a particular case, it is only a mnemonic.

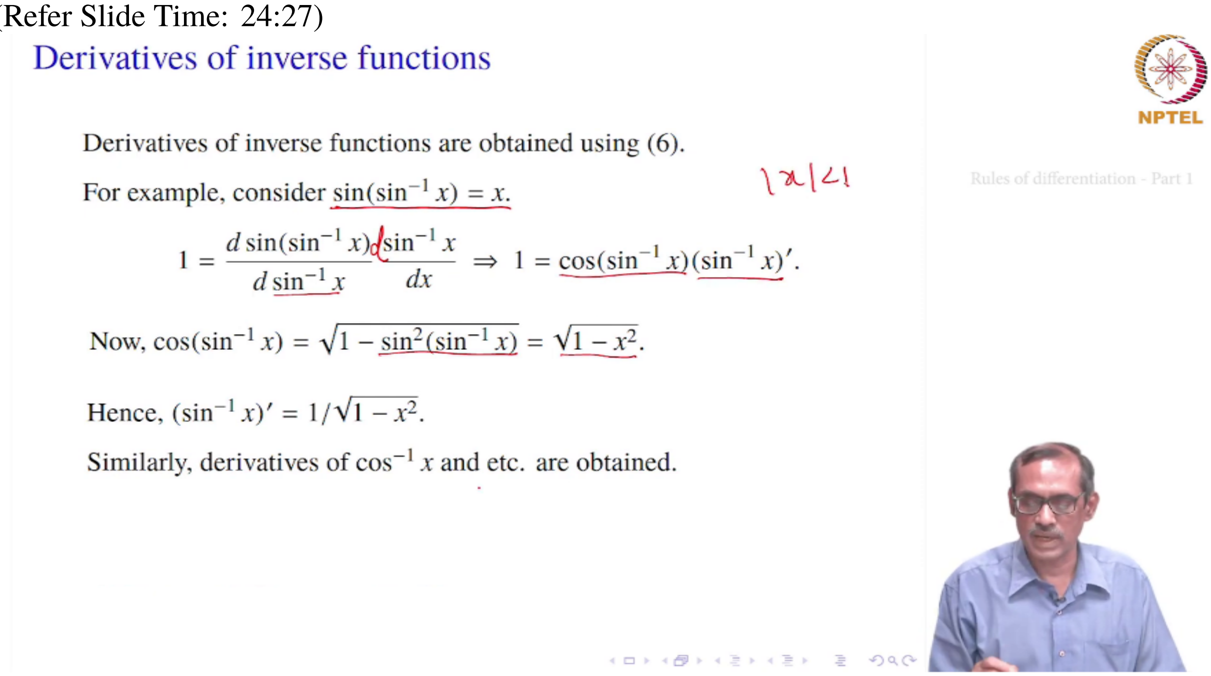
Similarly, we will get the derivative of $\sin x$ equal to $\cos x$. We have seen how to compute this in another problem. Of course, there you used the limit of $\sin(x/2)$ by $x/2$ as x goes to 0 is 1. Similarly, we have seen earlier that the derivative of $\cos x$ is $-\sin x$. You get the derivative of $\tan x$ as $\sec^x$. Of course, $\tan x$ is not defined at odd multiples of $\pi/2$. Except those points the derivative of $\tan x$ will be $\sec^2 x$.

Similarly, the derivative of $\sec x$ is $\sec x \tan x$. Again, except the points where it is not defined, everywhere else this formula holds. Then $\operatorname{cosec} x$ will give us $-\operatorname{cosec} x \cot x$. And, the derivative of $\cot x$ is equal to $-\operatorname{cosec}^2 x$. Again, we have to take care of where they are defined, where they are not. Then the derivative of $\sin^{-1} x$ will be $1/\sqrt{1-x^2}$. Remember, we define $\sin^{-1} x$ for $|x| < 1$. So, it gives us $1/\sqrt{1-x^2}$ for $|x| < 1$. Of course, $\sin^{-1} x$ is defined at $x = 1$, but it is not differentiable there. So, we will not have this.

We will see how to derive the derivative of this inverse functions. We will derive the formula for at least one of the inverses, other things are similar. It will come from the chain rule, as we have remarked earlier. (The rule number six). Similarly, if you take $\tan^{-1} x$, its derivative is $1/(1+x^2)$ for every x. And the derivative of $\sec^{-1} x$ is $1/(|x|\sqrt{x^2-1})$. When we go to $\cos^{-1} x$, it will be having a negative sign, because $\cos^{-1} x = \pi/2 - \sin^{-1} x$. So, 1 minus sin will come as $\pi/2$, whose derivative is 0. Similarly, if we take $\cot^{-1} x$, there will be a minus sign here, the derivative of $\operatorname{cosec}^{-1} x$ will have minus sign here. This is for $|x| > 1$.

Let us see how we have obtained the derivative of $\sin^{-1} x$. In fact, for all the inverse functions it will be obtained that way. For $\sin^{-1} x$, we consider this equation: $\sin(\sin^{-1} x) = x$. for $|x| < 1$. We are interested there only. Now we differentiate both sides of this and apply the chain rule. The right side gives $dx/dx = 1$. We keep that on the left side. The derivative of $\sin(\sin^{-1} x$ would give that function with respect to $\sin^{-1} x$ and $\sin^{-1} x$ with respect to x. This is like your $d \sin y/dy$ and dy/dx. That is d of sin function, that gives you $\cos y$. That is, $\cos(\sin^{-1} x)$. And we want to find out the derivative of $\sin^{-1} x$, one d is missing here, it is $d \sin^{-1} x/dx$, which is your $(\sin^{-1} x)'$.

(Refer Slide Time: 24:27)



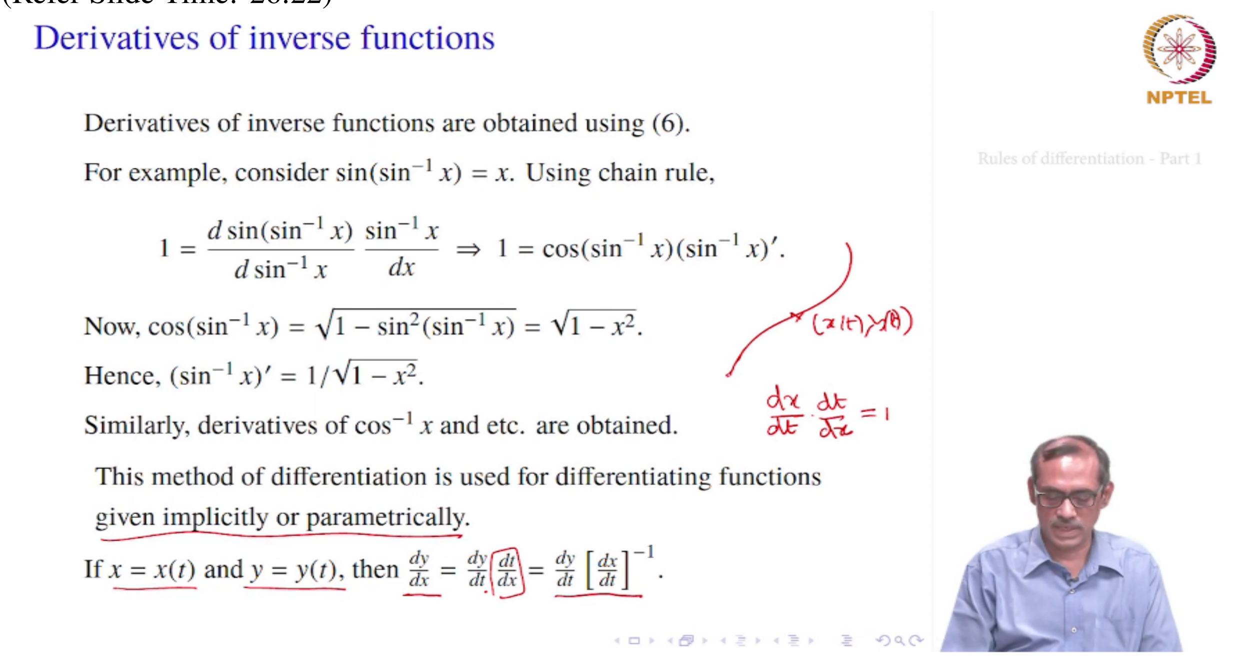
This gives $(\sin^{-1} x)'$, which is the reciprocal of $\cos(\sin^{-1} x)$. Now, what is $\cos(\sin^{-1} x)$? Imagine $\sin^{-1} x$ as t. So $\sqrt{1-\sin^2 t} = \cos t$ and $\sin(\sin^{-1} x) = x$ shows its square as x^2. So, $\cos(\sin^{-1} x) = \sqrt{1-x^2}$. That is the reason we get $\sqrt{1-x^2}$ here. And the derivative of $\sin^{-1} x$

will be the reciprocal of that. So, it is $1/\sqrt{1-x^2}$. That is how all the inverse functions can be differentiated.

In fact, sometimes we will not get the function given explicitly as $y = f(x)$. But there will be an equation, and we will say that the function is defined in its domain or in one of their possible domains. If it is implicit, it can be union of intervals. In one of those intervals we would want its differentiation. This equation can of course, be solved if possible, but we do not need to solve it for differentiation. We can differentiate the whole equation as we are doing for $\sin(\sin^{-1} x) = x$, and then simplify it to compute our derivative. That is called the implicit differentiation.

Sometimes you will be given x as a function of t and y as a function of t. The dependence of y on x is given through that t; that is called a parametric way of defining the function. So, we can also use our differentiation method if either it is given implicitly or it is given parametrically.

(Refer Slide Time: 26:22)



It may be like this one: $x = x(t)$, $y = y(t)$. It is some curve in the plane. This may not be a function $y = f(x)$, but the parameter t may be time or the length of the curve from the beginning. If that is your t, then every point can be thought of having coordinates $((x(t), y(t))$.

Suppose a function is given parametrically in terms of a parameter t. We want to find dy/dx. Again, we apply the chain rule. So, $dy/dx = (dy/dt)(dt/dx)$. Now, dt/dx is really $(dx/dt)^{-1}$. Again you can find that by differentiating the identity function. This one really is the identity function x. So, x is the function of t and again t is taken as a function of x; it is imagined that way. We then get the identity function $x = x(t(x))$. Once you differentiate by the chain rule it gives $(dx/dt)(dt/dx) = 1$. So, you get $dt/dx = (dx/dt)^{-1}$. We can use this formula now. It implies $dy/dx = (dy/dt)(dx/dt)^{-1}$. Of course, they will be defined when $dx/dt \neq 0$ at the point of concern. If we want abstractly for every t, then everywhere in the domain of the parameter, dx/dt should not be 0. This is assumed implicitly.