Introduction to Implicit Differentiation

We've learned a few specific and several general formulas for finding derivatives. Today we'll use the chain rule to further expand our ability to differentiate functions. Today's topic is implicit differentiation, which will allow us to differentiate many functions we haven't been able to differentiate yet.

Implicit Differentiation (Rational Exponent Rule)

We know that if n is an integer then the derivative of $y = x^n$ is nx^{n-1}. Does this formula still work if n is not an integer? I.e. is it true that:

$$\frac{d}{dx}(x^a) = ax^{a-1}.$$

We proved this formula using the definition of the derivative and the binomial theorem for $a = 1, 2, ...$ From this, we also got the formula for $a = -1, -2, ...$ Now we'll extend this formula to cover rational numbers $a = \frac{m}{n}$ as well. In particular, this will let us take the derivative of $y = \sqrt[n]{x} = x^{1/n}$.

Suppose $y = x^{\overline{n}}$ with m over the bar, i.e. $y = x^{\frac{m}{n}}$, where m and n are integers. We want to compute $\frac{dy}{dx}$. None of the rules we've learned so far seem helpful here, and if we use the definition of the derivative we'll get stuck trying to simplify $(x - \Delta x)^{m/n}$. We need a new idea.

The thing that's keeping us from using the definition of the derivative is that the denominator of n in the exponent forces us to take the n^{th} root of x. We could solve this problem by raising both sides of the equation to the n^{th} power:

$$\begin{aligned} y &= x^{\frac{m}{n}} \\ y^n &= (x^{\frac{m}{n}})^n \\ y^n &= x^{\frac{m}{n} \cdot n} \\ y^n &= x^m \end{aligned}$$

What happens if we try to take the derivative now by applying the operator $\frac{d}{dx}$? We have a rule for finding the derivative of a variable raised to an integer power; we can use this rule on both sides of the equation $y^n = x^m$.

$$\begin{aligned} y^n &= x^m \\ \frac{d}{dx}y^n &= \frac{d}{dx}x^m \end{aligned}$$

How do we compute $\frac{d}{dx}y^n$? We know that y is a function of x, so we can apply the chain rule with outside function y^n and inside function y. Suppose $u = y^n$. Then the chain rule tells us:

$$\frac{du}{dx} = \frac{du}{dy}\frac{dy}{dx}$$

So

$$\frac{d}{dx}y^n = \left(\frac{d}{dy}y^n\right)\frac{dy}{dx} = ny^{n-1}\frac{dy}{dx}.$$

On the right hand side of the equation we have $\frac{d}{dx}x^m = mx^{m-1}$, so we end up with:

$$\begin{aligned} \frac{d}{dx}y^n &= \frac{d}{dx}x^m \\ ny^{n-1}\frac{dy}{dx} = mx^{m-1} \end{aligned}$$

We're left with only one unknown quantity in this equation — $\frac{dy}{dx}$ — which is exactly what we're trying to find. Can we solve for $\dfrac{dy}{dx}$ and use this to find the derivative of $y = x^{m/n}$? We can, but we need to use a lot of algebra to do it.

By dividing both sides by ny^{n-1} we get:

$$\frac{dy}{dx} = \frac{m}{n}\frac{x^{m-1}}{y^{n-1}}$$

This looks promising but we want our answer in terms of x, without any y's mixed in. To get rid of the y we can now substitute $x^{m/n}$ for y. (We couldn't have done this before taking the derivative because we don't know how to take the derivative of $x^{m/n}$ — that's the whole point!)

$$\begin{aligned} \frac{dy}{dx} &= \frac{m}{n}\left(\frac{x^{m-1}}{y^{n-1}}\right) \\ &= \frac{m}{n}\left(\frac{x^{m-1}}{(x^{m/n})^{(n-1)}}\right) \\ &= \frac{m}{n}\left(\frac{x^{m-1}}{x^{(m/n)\cdot(n-1)}}\right) \\ &= \frac{m}{n}\frac{x^{m-1}}{x^{m(n-1)/n}} \\ &= \frac{m}{n}x^{((m-1)-\ \frac{m(n-1)}{n})} \\ &= \frac{m}{n}x^{\frac{n(m-1)}{n}\ -\ \frac{m(n-1)}{n}} \\ &= \frac{m}{n}x^{\frac{n(m-1)-m(n-1)}{n}} \\ &= \frac{m}{n}x^{\frac{nm-n-nm+m}{n}} \\ &= \frac{m}{n}x^{\frac{m-n}{n}} \\ &= \frac{m}{n}x^{(\frac{m}{n}\ -\ \frac{n}{n})} \\ \text{So,}\quad \frac{dy}{dx} &= \frac{m}{n}x^{(\frac{m}{n}\ -\ 1)} \end{aligned}$$

This is the answer we were hoping to get! We now know that for any rational number a, the derivative of x^a is ax^{a-1}.

Slope of a line tangent to a circle – direct version

A circle of radius 1 centered at the origin consists of all points (x, y) for which $x^2 + y^2 = 1$. This equation does not describe a function of x (i.e. it cannot be written in the form $y = f(x)$). Indeed, any vertical line drawn through the interior of the circle meets the circle in two points — every x has two corresponding y values. Let's see what goes wrong if we attempt to solve the equation of a circle for y in terms of x.

$$\begin{aligned} x^2 + y^2 &= 1 \\ x^2 + y^2 - x^2 &= 1 - x^2 \\ y^2 &= 1 - x^2 \\ y &= \pm\sqrt{1 - x^2} \end{aligned}$$

This still isn't a function because we get two choices for y — positive or negative. However, we do get a function if we look just at the positive case (i.e. at just the top half of the circle), and we can then find $\frac{dy}{dx}$, which will be the slope of a line tangent to the top half of the circle.

To compute this derivative, we first convert the square root into a fractional exponent so that we can use the rule from the previous example.

$$y = \sqrt{1 - x^2} = (1 - x^2)^{\frac{1}{2}}$$

Next, we need to use the chain rule to differentiate $y = (1 - x^2)^{\frac{1}{2}}$. The outside function is $u^{1/2}$ and the inside function is $1 - x^2$, so the chain rule tells us that

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}.$$

$$\frac{dy}{dx} = \frac{1}{2}u^{-1/2} \cdot (-2x) = -x \cdot (1 - x^2)^{-1/2} = \frac{-x}{\sqrt{1 - x^2}}.$$

If we want, we can use the fact that $y = \sqrt{1 - x^2}$ to rewrite this as $y' = -x/y$.

We conclude that the slope of the line tangent to a point (x, y) on the top half of the unit circle is $-x/y$.

Slope of a line tangent to a circle – implicit version

We just finished calculating the slope of the line tangent to a point (x, y) on the top half of the unit circle. In this calculation we started by solving the equation $x^2 + y^2 = 1$ for y, chose one "branch" of the solution to work with, then used the chain rule, the power rule and some algebra of exponents to compute the derivative $\frac{dy}{dx} = -\frac{x}{y}$.

We'll now see how we could have used implicit differentiation to do the same calculation much more easily. In fact, we'll find the slope of a line tangent to *any* point on the unit circle.

We don't need to solve for y — we can just apply the operator $\frac{d}{dx}$ to both sides of the original equation:

$$\begin{aligned} x^2 + y^2 &= 1 \\ \frac{d}{dx}\left(x^2 + y^2\right) &= \frac{d}{dx}(1) \\ \frac{d}{dx}(x^2) + \frac{d}{dx}(y^2) &= 0 \end{aligned}$$

We can easily take the derivative of the first term. For the second term, applying the chain rule with the inside function y and outside function u^2 gives us:

$$\begin{aligned} 2x + 2y\frac{dy}{dx} &= 0 \\ 2y\frac{dy}{dx} &= -2x \\ \frac{dy}{dx} &= -\frac{x}{y} \end{aligned}$$

This is the same answer, but we didn't have to restrict ourselves to just the top half of the circle or use any square roots. Implicit differentiation made this calculation much easier.

Implicit Differentiation Example

How would we find $y' = \frac{dy}{dx}$ if $y^4 + xy^2 - 2 = 0$?

We could use a trick to solve this explicitly — think of the above equation as a quadratic equation in the variable y^2 then apply the quadratic formula:

$$\begin{aligned} y^2 &= \frac{-x \pm \sqrt{x^2+8}}{2}, \\ &\text{so} \\ y &= \pm\sqrt{\frac{-x \pm \sqrt{x^2+8}}{2}}. \end{aligned}$$

Since we see $\pm$ twice in this equation, there are four possible branches to consider. This means that to be thorough we'd want to compute four different derivatives. This is a lot of work.

Instead, we can compute $\frac{dy}{dx}$ using implicit differentiation. As always, we start by applying $\frac{d}{dx}$ to both sides:

$$\begin{aligned} \frac{d}{dx}(y^4 + xy^2 - 2) &= \frac{d}{dx}0 \\ \frac{d}{dx}(y^4) + \frac{d}{dx}(xy^2) - \frac{d}{dx}2 &= 0 \\ 4y^3\frac{dy}{dx} + (y^2 + x \cdot 2y\frac{dy}{dx}) - 0 &= 0 \\ 4y^3\frac{dy}{dx} + 2xy\frac{dy}{dx} &= -y^2 \\ (4y^3 + 2xy)\frac{dy}{dx} &= -y^2 \\ \frac{dy}{dx} &= \frac{-y^2}{4y^3 + 2xy} \end{aligned}$$

In lecture Professor Jerison used the shorthand y' for the derivative; here we use $\frac{dy}{dx}$ to make it clear that we are differentiating with respect to x.

Derivative of the Inverse of a Function

One very important application of implicit differentiation is to finding derivatives of inverse functions.

We start with a simple example. We might simplify the equation $y = \sqrt{x}$ $(x > 0)$ by squaring both sides to get $y^2 = x$. We could use function notation here to say that $y = f(x) = \sqrt{x}$ and $x = g(y) = y^2$.

In general, we look for functions $y = f(x)$ and $g(y) = x$ for which $g(f(x)) = x$. If this is the case, then g is the inverse of f (we write $g = f^{-1}$) and f is the inverse of g (we write $f = g^{-1}$).

How are the graphs of a function and its inverse related? We start by graphing $f(x) = \sqrt{x}$. Next we want to graph the inverse of f, which is $g(y) = x$. But this is exactly the graph we just drew. To compare the graphs of the functions f and f^{-1} we have to exchange x and y in the equation for f^{-1}. So to compare $f(x) = \sqrt{x}$ to its inverse we replace y's by x's and graph $g(x) = x^2$.

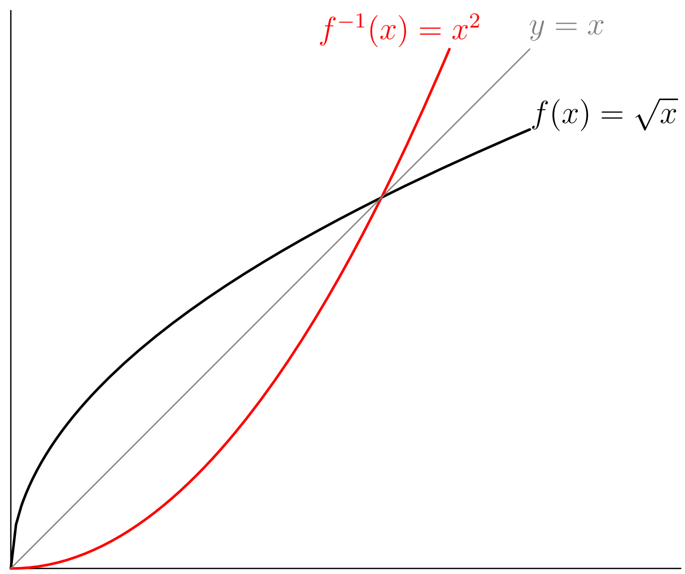

Figure 1: The graph of f^{-1} is the reflection of the graph of f across the line $y = x$

In general, if you have the graph of a function f you can find the graph of f^{-1} by exchanging the x- and y-coordinates of all the points on the graph. In other words, the graph of f^{-1} is the reflection of the graph of f across the line $y = x$.

This suggests that if $\frac{dy}{dx}$ is the slope of a line tangent to the graph of f, then

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}$$

is the slope of a line tangent to the graph of f^{-1}. We could use the definition of the derivative and properties of inverse functions to turn this suggestion into a proof, but it's easier to prove using implicit differentiation.

Let's use implicit differentiation to find the derivative of the inverse function:

$$\begin{aligned} y &= f(x) \\ f^{-1}(y) &= x \\ \frac{d}{dx}(f^{-1}(y)) &= \frac{d}{dx}(x) = 1 \end{aligned}$$

By the chain rule:

$$\frac{d}{dy}(f^{-1}(y))\frac{dy}{dx} = 1$$

so

$$\frac{d}{dy}(f^{-1}(y)) = \frac{1}{\frac{dy}{dx}}.$$

Implicit differentiation allows us to find the derivative of the inverse function $x = f^{-1}(y)$ whenever we know the derivative of the original function $y = f(x)$.

Derivative of $\arctan(x)$

Let's use our formula for the derivative of an inverse function to find the derivative of the inverse of the tangent function: $y = \tan^{-1} x = \arctan x$.

We simplify the equation by taking the tangent of both sides:

$$\begin{aligned} y &= \tan^{-1} x \\ \tan y &= \tan(\tan^{-1} x) \\ \tan y &= x \end{aligned}$$

To get an idea what to expect, we start by graphing the tangent function (see Figure 1). The function $\tan(x)$ is defined for $-\frac{\pi}{2} < x < \frac{\pi}{2}$. It's graph extends from negative infinity to positive infinity.

If we reflect the graph of $\tan x$ across the line $y = x$ we get the graph of $y = \arctan x$ (Figure 2). Note that the function $\arctan x$ is defined for all values of x from minus infinity to infinity, and $\lim_{x\to\infty} \tan^{-1} x = \frac{\pi}{2}$.

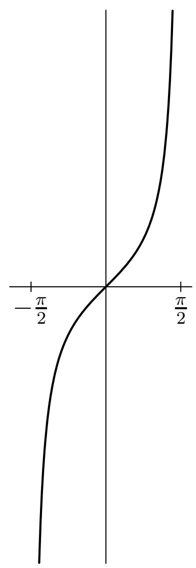

Figure 1: Graph of the tangent function.

You may know that:

$$\begin{aligned} \frac{d}{dy}\tan y &= \frac{d}{dy}\frac{\sin y}{\cos y} \\ &\vdots \\ &= \frac{1}{\cos^2 y} \\ &= \sec^2 y \end{aligned}$$

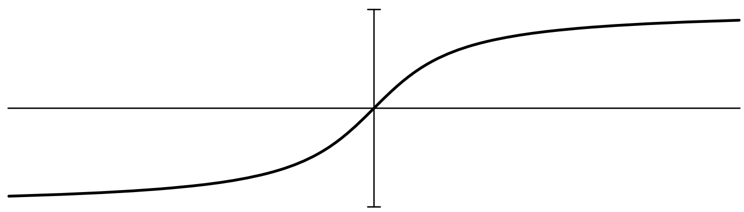

Figure 2: Graph of $\tan^{-1} x$.

(If you haven't seen this before, it's good exercise to use the quotient rule to verify it!)

We can now use implicit differentiation to take the derivative of both sides of our original equation to get:

$$\begin{aligned} \tan y &= x \\ \frac{d}{dx}(\tan(y)) &= \frac{d}{dx}x \\ \text{(Chain Rule)} \quad \frac{d}{dy}(\tan(y))\frac{dy}{dx} &= 1 \\ \left(\frac{1}{\cos^2(y)}\right)\frac{dy}{dx} &= 1 \\ \frac{dy}{dx} &= \cos^2(y) \end{aligned}$$

Or equivalently, $y' = \cos^2 y$. Unfortunately, we want the derivative as a function of x, not of y. We must now plug in the original formula for y, which was $y = \tan^{-1} x$, to get $y' = \cos^2(\arctan(x))$. This is a correct answer but it can be simplified tremendously. We'll use some geometry to simplify it.

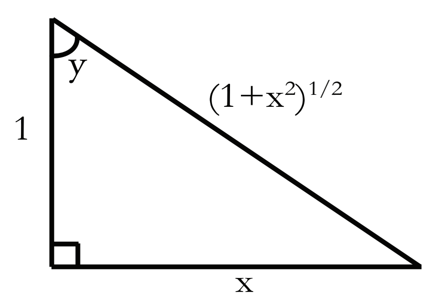

Figure 3: Triangle with angles and lengths corresponding to those in the example.

In this triangle, $\tan(y) = x$ so $y = \arctan(x)$. The Pythagorean theorem tells us the length of the hypotenuse:

$$h = \sqrt{1 + x^2}$$

and we can now compute:

$$\cos(y) = \frac{1}{\sqrt{1 + x^2}}.$$

From this, we get:

$$\cos^2(y) = \left(\frac{1}{\sqrt{1 + x^2}}\right)^2 = \frac{1}{1 + x^2}$$

so:

$$\frac{dy}{dx} = \frac{1}{1 + x^2}.$$

In other words,

$$\frac{d}{dx}\arctan(x) = \frac{1}{1 + x^2}.$$

Derivative of $\arcsin(x)$

For a final example, we quickly find the derivative of $y = \sin^{-1} x = \arcsin x$.

As usual, we simplify the equation by taking the sine of both sides:

$$\begin{aligned} y &= \sin^{-1} x \\ \sin y &= x \end{aligned}$$

We next take the derivative of both sides of the equation and solve for $y' = \frac{dy}{dx}$.

$$\begin{aligned} \sin y &= x \\ (\cos y) \cdot y' &= 1 \\ y' &= \frac{1}{\cos y} \end{aligned}$$

We want to rewrite this in terms of $x = \sin y$. Luckily there is a simple trig. identity relating $\cos y$ to $\sin y$. We can solve it for $\cos y$ and "plug in".

$$\begin{aligned} \cos^2 y + \sin^2 y &= 1 \\ \cos^2 y &= 1 - \sin^2 y \\ \cos y &= \sqrt{1 - \sin^2 y} \quad (\cos y > 0 \text{ on the range of } y = \sin^{-1} x) \end{aligned}$$

Plugging this in to our equation for $y' = \frac{d}{dx} \sin^{-1} x$ we get:

$$y' = \frac{1}{\cos y} = \frac{1}{\sqrt{1 - \sin^2 y}} = \frac{1}{\sqrt{1 - x^2}}.$$

Notice that we made a choice between a positive and negative square root when solving for $\cos y$. We chose the positive square root because we usually define $\sin^{-1} x$ to have outputs between $-\pi/2$ and $\pi/2$, and the cosine function is always positive on this interval.

When dealing with inverse functions we are often faced with choices like this; when in doubt draw a graph and be sure your choices make sense in the context of your problem.

Differentiating Logs and Exponentials

We finally discuss derivatives of exponential and logarithmic functions. These are probably the only functions you're aware of that you're still unable to differentiate. Logs and exponentials are as fundamental as trigonometric functions, if not more so.

Working with exponents

We start out with "base" number a. This number a must be positive, and we're going to assume $a > 1$ to make it easier to draw graphs.

What is the derivative of a^x? We'll start to answer this question by reviewing what we know about exponents.

To begin with, we know that:

$$a^0 = 1; \quad a^1 = a; \quad a^2 = a \cdot a; \quad a^3 = a \cdot a \cdot a \quad \dots$$

In general,

$$a^{x_1 + x_2} = a^{x_1} a^{x_2}$$

Together with the first two properties, this describes the exponential function a^x.

From these properties, we can derive:

$$(a^{x_1})^{x_2} = a^{x_1 x_2}$$

and we can easily evaluate a^n for any positive integer n. For negative integers, we can see from the fact that $a^m \cdot a^{-m} = a^{m-m} = 1$ that $a^{-m} = \frac{1}{a^m}$.

We want to be able to evaluate a^x for any number x; not just for integers. We start by defining a^x for rational values of x:

$$a^{\frac{p}{q}} = \sqrt[q]{a^p} \quad \text{(where } p \text{ and } q \text{ are integers.)}$$

Since $a^{1/2} \cdot a^{1/2} = a^1 = \sqrt{a} \cdot \sqrt{a}$, this seems like a reasonable definition.

All that's left is to define a^x for irrational numbers; we do this by "filling in" the gaps in the function to make it continuous. This is what your calculator does when you ask it for the value of $3^{\sqrt{2}}$ or 2^{π}. It can't give you an exact answer, so it gives you a decimal (rational) number very close to the exact answer.

Take some time and sketch the graph of 2^x to "get a feel" for how exponential functions work.

a^x and the Definition of the Derivative

Our goal is to calculate the derivative $\frac{d}{dx}a^x$. It's going to take us a while.

We start by writing down the definition of the derivative

$$\frac{d}{dx}a^x = \lim_{\Delta x \to 0} \frac{a^{x+\Delta x} - a^x}{\Delta x}$$

We can use the rule $a^{x_1+x_2} = a^{x_1}a^{x_2}$ to factor out a^x:

$$\begin{aligned}
\frac{d}{dx}a^x &= \lim_{\Delta x \to 0} \frac{a^{x+\Delta x} - a^x}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{a^x a^{\Delta x} - a^x}{\Delta x} \\
&= \lim_{\Delta x \to 0} a^x \frac{a^{\Delta x} - 1}{\Delta x}
\end{aligned}$$

As we're taking this limit, we're holding a and x fixed while Δx changes (approaches zero). This means that for the purposes of taking this limit, a^x is a constant. We can therefore factor the constant multiple out of the limit to get:

$$\frac{d}{dx}a^x = a^x \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}$$

We've made a good start at finding the derivative of a^x; let's look at what we have so far. We can see from our calculations that $\frac{d}{dx}a^x$ is a^x times some multiple whose value we don't yet know. Let's call that multiple $M(a)$:

$$M(a) = \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}.$$

Using this definition of $M(a)$, we can say that $\frac{d}{dx}a^x = M(a)a^x$.

Slope of the tangent to a^x

We defined a function $M(a)$ as follows:

$$M(a) = \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}.$$

This definition allows us to say that $\frac{d}{dx} a^x = M(a)a^x$. In order to understand the derivative of a^x we must understand $M(a)$; we next look at two different ways of thinking about $M(a)$.

First, if we plug $x = 0$ in to the definition of the derivative of a^x we get:

$$\begin{aligned}
\left.\frac{d}{dx} a^x\right|_{x=0} &= \left.\lim_{\Delta x \to 0} \frac{a^{x+\Delta x} - a^x}{\Delta x}\right|_{x=0} \\
&= \lim_{\Delta x \to 0} \frac{a^{0+\Delta x} - a^0}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x} \\
&= M(a)
\end{aligned}$$

(or we could simply observe that $\frac{d}{dx} a^x|_{x=0} = M(a)a^0 = M(a)$). So $M(a)$ is the value of the derivative of a^x when $x = 0$.

Remember that the derivative tells us the slope of the tangent line to the graph. So $M(a)$ can also be thought of as the slope of the graph of $y = a^x$ at $x = 0$.

Note that the shape of the graph of a^x depends on the choice of a, so for different values for a we'll get different tangent lines and different values for $M(a)$.

Because $\frac{d}{dx} a^x = M(a)a^x$, we only need to know the slope of the line tangent to the graph at $x = 0$ in order to figure out the slope of the tangent line at any point on the graph!

Remember that when we computed the derivative of the sine function we worked hard to compute the value of $\lim_{x \to 0} \frac{\sin x}{x}$. This value is just the derivative of $\sin x$ when $x = 0$ — check this yourself by writing down the definition of the derivative of $\sin x$ and replacing x by 0. In order to get a general formula for the derivative of the sine function we first had to know the value of its derivative when $x = 0$.

The formula for $a^{x+\Delta x}$ is simpler than the one for $\sin(x + \Delta x)$, so the first part of our calculation of $\frac{d}{dx} a^x$ was easier than the corresponding calculation for $\sin x$. But when we try to compute:

$$M(a) = \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}$$

we get stuck. We were able to use radians and the unit circle to find $\lim_{x \to 0} \frac{\sin x}{x}$,

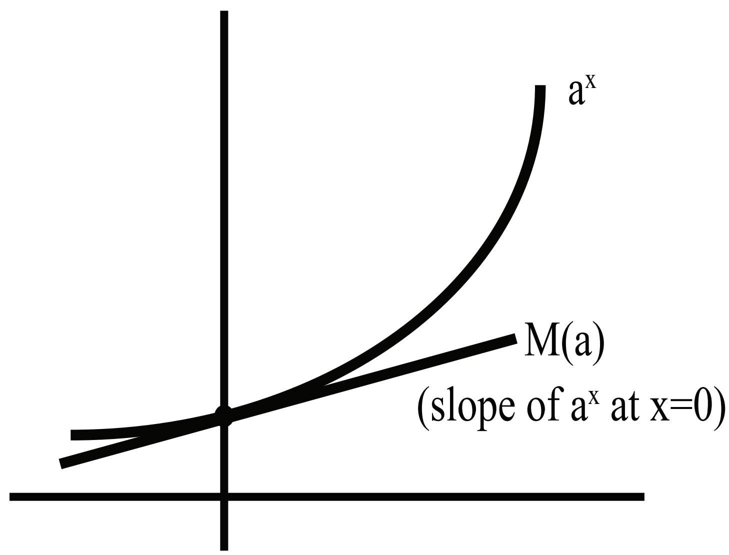

Figure 1: Geometric definition of $M(a)$

but we don't have a good way to find the exact slope of the tangent line to $y = a^x$ at $x = 0$.

Definition of e

Recall that:

$$M(a) = \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}.$$

is the value for which $\frac{d}{dx}a^x = M(a)a^x$, the value of the derivative of a^x when $x = 0$, and the slope of the graph of $y = a^x$ at $x = 0$. We need to know what $M(a)$ is in order to find out what the derivative of a^x is. It turns out that the easiest way to understand $M(a)$ is to give up trying to calculate it and to *define* e as the number such that $M(e) = 1$.

Leaving aside the question of whether such a number e exists, let's discuss what such a number would do for us. Since $M(e) = 1$,

$$\frac{d}{dx}e^x = e^x.$$

This is an incredibly important formula and is the only thing we've said so far this lecture that you need to memorize. Also, the slope of the tangent line to $y = e^x$ at $x = 0$ has slope 1. You can confirm this by plugging $x = 0$ into $\frac{d}{dx}e^x = e^x$.

But we still don't know what e is, or even if there is such a number. How do we know that there is *any* number a for which the slope of the tangent line to $y = a^x$ is 1 when $x = 0$?

First notice that as the base a increases, the graph of $y = a^x$ gets steeper. Is the slope ever 1?

If $a = 1$, $a^x = 1$ for all x and the slope of the tangent line to the (very simple) graph at $x = 0$ is 0. Although we may not be able to compute the slope exactly, we can use secant lines to estimate the slope $M(a)$ for $a = 2$ and $a = 4$ geometrically. Look at the graph of 2^x in Fig. 1. The secant line from $(0, 1)$ to $(1, 2)$ of the graph $y = 2^x$ has slope 1. We can see from the picture that the slope of $y = 2^x$ at $x = 0$ is less than the slope of this secant line: $M(2) < 1$ (see Fig. 1).

Next, look at the graph of 4^x in Fig. 2. The secant line from $(-\frac{1}{2}, \frac{1}{2})$ to $(1, 0)$ on the graph of $y = 4^x$ has slope 1. We see that the slope of $y = 4^x$ at $x = 0$ is greater than the slope of the secant, so $M(4) > 1$ (see Fig. 2).

Assuming our function M is continuous, we conclude that somewhere in between 2 and 4 there is a base whose slope at $x = 0$ is 1.

Thus we can *define* e to be the unique number such that

$$M(e) = 1$$

or, to put it another way,

$$\lim_{h \to 0} \frac{e^h - 1}{h} = 1$$

or, to put it yet another way,

$$\frac{d}{dx}(e^x) = 1 \quad \text{at } x = 0$$

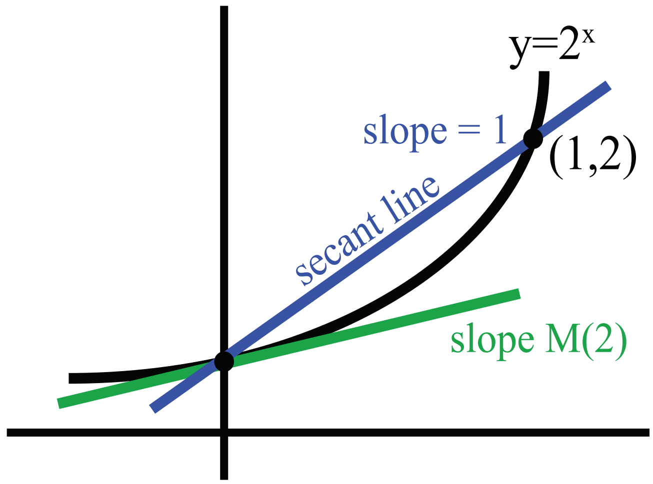

Figure 1: Slope $M(2) < 1$

Another way to convince ourselves that e must exist is to start with the graph of $f(x) = 2^x$ (recalling that $M(2) < 1$) and think about the function $f(kx) = 2^{kx}$. As k increases, the graph of $y = f(kx)$ is compressed horizontally and the slope of the tangent line to the graph of $y = f(x)$ continuously grows steeper. So, for some value of k between 1 and 2, the slope of that tangent line must be 1. So e exists and is between $2^1 = 2$ and $2^2 = 4$.

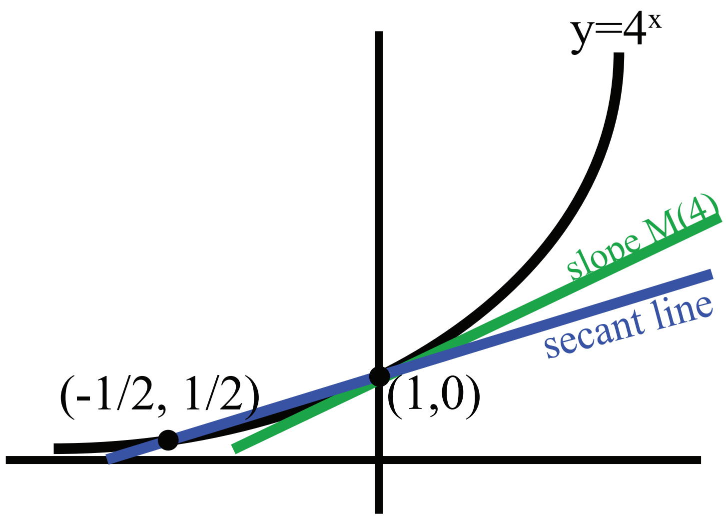

Figure 2: Slope $M(4) > 1$

Natural log (inverse function of e^x)

Recall that:

$$M(a) = \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}.$$

is the value for which $\frac{d}{dx} a^x = M(a) a^x$, the value of the derivative of a^x when $x = 0$, and the slope of the graph of $y = a^x$ at $x = 0$. To understand $M(a)$ better, we study the natural log function $\ln(x)$, which is the inverse of the function e^x. This function is defined as follows:

$$\text{If } \ y = e^x, \ \text{ then } \ \ln(y) = x$$

or

$$\text{If } \ w = \ln(x), \ \text{ then } \ e^w = x$$

Before we go any further, let's review some properties of this function:

$$\ln(x_1 x_2) = \ln x_1 + \ln x_2$$

$$\ln 1 = 0$$

$$\ln e = 1$$

These can be derived from the definition of $\ln x$ as the inverse of the function e^x, the definition of e, and the rules of exponents we reviewed at the start of lecture.

We can also figure out what the graph of $\ln x$ must look like. We know roughly what the graph of e^x looks like, and the graph of $\ln x$ is just the reflection of that graph across the line $y = x$. Try sketching the graph of $\ln x$ yourself.

You should notice the following important facts about the graph of $\ln x$. Since e^x is always positive, the domain (set of possible inputs) of $\ln x$ includes only the positive numbers. The entire graph of $\ln x$ lies to the right of the y-axis. Since $e^0 = 1$, $\ln 1 = 0$ and the graph of $\ln x$ goes through the point $(1, 0)$. And finally, since the slope of the tangent line to $y = e^x$ is 1 where the graph crosses the y-axis, the slope of the graph of $y = \ln x$ must be $1/1 = 1$ where the graph crosses the x-axis.

We know that $\frac{d}{dx} e^x = e^x$. To find $\frac{d}{dx} \ln x$ we'll use implicit differentiation as we did in previous lectures.

We start with $w = \ln(x)$ and compute $\dfrac{dw}{dx} = \dfrac{d}{dx} \ln x$. We don't have a good way to do this directly, but since $w = \ln(x)$, we know $e^w = e^{\ln(x)} = x$. We now use implicit differentiation to take the derivative of both sides of this equation.

$$\begin{aligned} \frac{d}{dx}(e^w) &= \frac{d}{dx}(x) \\ \frac{d}{dw}(e^w)\frac{dw}{dx} &= 1 \end{aligned}$$

$$\begin{aligned} e^w \frac{dw}{dx} &= 1 \\ \frac{dw}{dx} &= \frac{1}{e^w} = \frac{1}{x} \end{aligned}$$

So

$$\frac{d}{dx}(\ln(x)) = \frac{1}{x}$$

This is another formula worth memorizing.

$\frac{d}{dx}a^x$?

We now want to learn to differentiate *any* exponential a^x. There are two roughly equivalent methods we can use:

Method 1: Convert a^x to something with base e and use the chain rule.

Because $\ln x$ is the inverse function to e^x we can rewrite a as $e^{\ln(a)}$. Thus:

$$a^x = \left(e^{\ln(a)}\right)^x = e^{x\ln(a)}$$

That looks like it might be tricky to differentiate. Let's work up to it:

$$\begin{aligned}\frac{d}{dx}e^x &= e^x \\ \frac{d}{dx}e^{3x} &= 3e^{3x} \quad \text{(by the chain rule)}\end{aligned}$$

Remember, $\ln(a)$ is just a constant like 3, not a variable. Therefore:

$$\frac{d}{dx}e^{(\ln a)x} = (\ln a)e^{(\ln a)x}$$

or

$$\frac{d}{dx}a^x = (\ln a)a^x$$

This is a common type of calculation; you should practice it until you are comfortable with it. You may either memorize formulas for $\frac{d}{dx}e^{kx}$ and $\frac{d}{dx}a^x$ or re-derive them every time you need them.

Recall that $\frac{d}{dx}a^x = M(a) \cdot a^x$. So finally we know the value of $M(a)$:

$$M(a) = \ln(a)$$

The Most Natural Logarithmic Function

At times in your life you might find yourself tempted to use logarithmic functions with bases other than e; for example, $\log_2 n$ or $\log_{10} r$. We claim that $\ln x$, the natural logarithm or log base e, is the most natural choice of logarithmic function.

Today's example is from the field of economics. Imagine that the price of a stock you own goes down by a dollar; how does that affect you?

That depends on a lot of things. In particular, it depends on whether the original stock price was a dollar or a hundred dollars.

As another example, at the time this lecture was given the London Exchange FTSE index had closed down 27.9 points. This statistic is almost meaningless unless you know the actual total of the index, which was 6,432 at the time.

A more meaningful statistic is the change in the price divided by the price:

$$\frac{\Delta p}{p} = \frac{27.9}{6432} \approx 0.43\%.$$

This tells us that the FTSE index dropped by 0.43% of its total value today.

A day is a relatively short time in the life of an investment so if price p is a function of time t with $\Delta t = 1$ day,

$$\frac{\Delta p}{1 \text{ day}} \approx \frac{dp}{dt} = p'.$$

So instead of looking at $\frac{\Delta p}{p}$ we could discuss $\frac{p'}{p}$, which is just the derivative of the natural log of p.

$$\frac{p'}{p} = (\ln p)'$$

This is the formula for logarithmic differentiation, and it is used all the time by economists and people who model prices of things.

There's no point in using log base ten or log base two, because when you take the derivative of those functions you get an extra constant factor in the denominator:

$$(\log_{10} p)' = \frac{p'}{p \ln 10}.$$

This is just one example — any variable that has to do with ratios is going to involve logarithms. We'll see more of this when we study applications of derivatives. You may have guessed this already; the derivative of $\ln x$ is quite elegant and simple. The most elegant and simple ideas that are often the most powerful, as well.

The Functions 10^x and 2^x

We computed that $\frac{d}{dx}a^x = (\ln a)a^x$.

So

$$\frac{d}{dx}2^x = (\ln 2)2^x$$

and

$$\frac{d}{dx}10^x = (\ln 10)10^x.$$

Even if we insist on starting with another base, like 2 or 10, the natural logarithm appears. They come up naturally, independent of our human preferences like base 2 or base 10. The base e may seem strange at first, but it comes up everywhere. After a while you'll learn to appreciate just how natural it is.

$\dfrac{d}{dx}a^x$, part 2

We're learning to differentiate *any* exponential a^x. This is the second of two possible methods.

Method 2: Logarithmic differentiation

It turns out that sometimes it is hard to differentiate a function u and easier to differentiate $\ln u$ (for example, $u = e^{x^2+6}$.) We'd like to be able to use $\frac{d}{dx}\ln u$ to find $\frac{d}{dx}u$.

The chain rule tells us that $\frac{d}{dx}\ln u = \frac{d\ln u}{du}\frac{du}{dx}$, and we know that $\frac{d}{du}\ln u = \frac{1}{u}\frac{du}{dx}$, so

$$(\ln u)' = u'/u.$$

How does this help us compute $\frac{d}{dx}a^x$?

$$\begin{aligned} u &= a^x \\ \ln u &= \ln(a^x) \\ \ln u &= x\ln a \end{aligned}$$

This is pretty easy to differentiate because $\ln a$ is a constant:

$$(\ln u)' = \ln a.$$

Since $(\ln u)' = u'/u$, $u' = u(\ln u)'$. So $\frac{d}{dx}a^x = a^x \ln a = (\ln a)a^x$.

This uses the same arithmetic as the first method, but we don't have to convert to base e.

Another Example of Logarithmic Differentiation

This example could be done equally well by converting to base e, but we're going to do it using logarithmic differentiation. Recall that the rule we use for logarithmic differentiation is $(\ln u)' = u'/u$.

Here we have a "moving" (non-constant) exponent and a moving base.

Example: Let $v = x^x$. Find v'.

First, we take the natural log of both sides to see that $\ln v = \ln(x^x) = x \ln x$.

Next, we differentiate both sides of the equation, using the product rule and the rule for the derivative of $\ln x$ on the right hand side:

$$(\ln v)' = \ln x + x \cdot \frac{1}{x}.$$

Now apply the formula $(\ln u)' = u'/u$. to get:

$$v'/v = 1 + \ln x$$

Plugging in x^x for v and solving for v', we get:

$$\begin{aligned} \frac{v'}{x^x} &= 1 + \ln x \\ v' &= x^x(1 + \ln x) \\ \frac{d}{dx}x^x &= x^x(1 + \ln x) \end{aligned}$$

The Power Rule

What is the derivative of $\frac{d}{dx}x^r$? We answered this question first for positive integer values of r, for all integers, and then for rational values of r:

$$\frac{d}{dx}x^r = rx^{r-1}$$

We'll now prove that this is true for any *real* number r. We can do this two ways:

1st method: base e

Since $x = e^{\ln x}$, we can say:

$$\begin{aligned} x^r &= \left(e^{\ln x}\right)^r \\ x^r &= e^{r \ln x} \end{aligned}$$

We take the derivative of both sides to get:

$$\begin{aligned} \frac{d}{dx}x^r &= \frac{d}{dx}e^{r\ln x} = e^{r\ln x}\frac{d}{dx}(r\ln x) \quad \text{(by the chain rule)} \\ &= e^{r\ln x}\left(\frac{r}{x}\right) \quad \text{(remember } r \text{ is constant)} \\ &= x^r\left(\frac{r}{x}\right) \quad \text{(because } x^r = e^{r\ln x}\text{)} \\ \frac{d}{dx}x^r &= rx^{r-1} \end{aligned}$$

2nd method: logarithmic differentiation

We define $f(x) = x^r$, and take the natural log of both sides to get $\ln f = r \ln x$. The technique of logarithmic differentiation requires us to we plug our function into the formula:

$$(\ln f)' = \frac{f'}{f}$$

So we first compute:

$$\begin{aligned} \ln f &= \ln x^r \\ \ln f &= r \ln x \end{aligned}$$

And then take the derivative of both sides to get:

$$(\ln f)' = \frac{r}{x}$$

Since $(\ln f)' = \frac{f'}{f}$, we have:

$$f' = f(\ln f)' \quad = \quad x^r \left(\frac{r}{x}\right) = rx^{r-1}.$$

Look over the two methods again – the calculations are almost the same. This is typical. To use the second method we had to introduce a new symbol like u or f. In the first method we had to deal with exponents. It's worthwhile know both methods.

Another Moving Exponent

Find the value of:

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

Technically this is not a calculus problem, but we will use some calculus to solve it. There are two reasons to discuss this now – first that the answer is very interesting and second that it has another moving exponent – the exponent n in the problem is changing as we take the limit.

Whenever we're faced with a moving exponent our first step is to use a logarithm to turn the exponent into a multiple:

$$\ln\left[\left(1+\frac{1}{n}\right)^n\right] = n\ln\left(1+\frac{1}{n}\right).$$

Now we want to start thinking about the limit of this quantity as n approaches infinity. We've had a lot of practice thinking about limits as Δx approaches zero and very little practice with numbers approaching infinity, so it makes sense to try to rephrase this from a question about a very large number n to a question about a very small number Δx.

The quantity $\Delta x = 1/n$ will approach zero as n goes to infinity. If $\Delta x = 1/n$ then $n = 1/\Delta x$, and we get:

$$\lim_{n\to\infty}\left[n\ln\left(1+\frac{1}{n}\right)\right] = \lim_{\Delta x\to 0}\left[\frac{1}{\Delta x}\ln(1+\Delta x)\right].$$

This doesn't look like much of an improvement, but by subtracting $0 = \ln 1$ from $\ln(1+\Delta x)$ we can put it in a familiar form:

$$\begin{aligned}
\lim_{\Delta x\to 0}\left[\frac{1}{\Delta x}\ln(1+\Delta x)\right] &= \lim_{\Delta x\to 0}\left[\frac{1}{\Delta x}\left(\ln(1+\Delta x)-\ln 1\right)\right] \\
&= \lim_{\Delta x\to 0}\frac{\ln(1+\Delta x)-\ln 1}{\Delta x} \\
&= \left.\frac{d}{dx}\ln x\right|_{x=1} \\
&= \left.\frac{1}{x}\right|_{x=1} \\
&= 1
\end{aligned}$$

By strategically subtracting zero ($\ln 1$), we were able to turn this ugly limit into a difference equation, which we could then evaluate using calculus.

Now we just have to work backward to figure out the answer to our original question.

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \lim_{n\to\infty} e^{\ln\left[\left(1+\frac{1}{n}\right)^n\right]}$$

$$
\begin{aligned}
&= e^{\lim_{n\to\infty} \ln\left[\left(1+\frac{1}{n}\right)^n\right]} \\
&= e^1 \\
&= e
\end{aligned}
$$

That's right,

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$$

and we now have a way to get a numerical value for e. Using this formula we can find the value of e with as much precision as our calculators will allow. For example,

$$\left(1+\frac{1}{10000}\right)^{10000} \cong 2.7182.$$

A Formula for e

We calculated that if $a_k = (1 + \frac{1}{k})^k$, then $\lim_{k\to\infty} a_k = e$ by first showing that $\lim_{k\to\infty} \ln a_k = 1$. Since $e^{\ln a_k} = a_k$, as k goes to infinity $a_k = e^{\ln a_k}$ will tend toward $e^1 = e$.

The key point here was that $a_k = e^{\ln a_k}$; that the natural log function is the inverse of the exponential function.

Question: Shouldn't $\ln a_k$ tend towards zero, because a_k tends toward 1?

Answer: It's true that $(1+\frac{1}{k})$ tends toward 1, and so $\ln(1+\frac{1}{k})$ tends toward 0. But that's not the limit we want; we're asking about $\ln a_k = k \cdot \ln(1 + \frac{1}{k})$. As $\ln(1 + \frac{1}{k})$ is tending toward 0, k tends toward infinity. That's why we needed to use limits and derivatives to figure out what the limit of this expression was.

We know that $\lim_{k\to\infty} a_k = e$, and all equalities can be read in both directions. So $e = \lim_{k\to\infty} \left(1 + \frac{1}{k}\right)^k$. In other words, this limit is a formula for e. By looking at the formula from a different angle, we discover that we can use the expression $\left(1 + \frac{1}{k}\right)^k$ to compute a base e for which graph of e^x has slope 1 when $x = 0$.

Often we can improve our understanding of mathematics by looking at things in several different ways, and that's what we're going to be doing at the end of this lecture on derivatives.

Derivatives of Hyperbolic Sine and Cosine

Hyperbolic sine (pronounced "sinsh"):

$$\sinh(x) = \frac{e^x - e^{-x}}{2}$$

Hyperbolic cosine (pronounced "cosh"):

$$\cosh(x) = \frac{e^x + e^{-x}}{2}$$

$$\frac{d}{dx}\sinh(x) = \frac{d}{dx}\left(\frac{e^x - e^{-x}}{2}\right) = \frac{e^x - (-e^{-x})}{2} = \cosh(x)$$

Likewise,

$$\frac{d}{dx}\cosh(x) = \sinh(x)$$

(Note that this is different from $\frac{d}{dx}\cos(x)$.)

Important identity:

$$\cosh^2(x) - \sinh^2(x) = 1$$

Proof:

$$\begin{aligned}
\cosh^2(x) - \sinh^2(x) &= \left(\frac{e^x + e^{-x}}{2}\right)^2 - \left(\frac{e^x - e^{-x}}{2}\right)^2 \\
\cosh^2(x) - \sinh^2(x) &= \frac{1}{4}\left(e^{2x} + 2e^x e^{-x} + e^{-2x}\right) - \frac{1}{4}\left(e^{2x} - 2 + e^{-2x}\right) = \frac{1}{4}(2+2) = 1
\end{aligned}$$

Why are these functions called "hyperbolic"?
Let $u = \cosh(x)$ and $v = \sinh(x)$, then

$$u^2 - v^2 = 1$$

which is the equation of a hyperbola.

Regular trig functions are "circular" functions. If $u = \cos(x)$ and $v = \sin(x)$, then

$$u^2 + v^2 = 1$$

which is the equation of a circle.