Welcome to 18.01

Welcome to 18.01. Today we start "Unit One"; the topic of the unit is differentiation. We'll start by reviewing what's in store in the next couple of weeks.

The topic of this lecture is "what is a derivative?" We're going to look at this question from several different points of view, and the first one is the geometric interpretation. We'll also discuss a physical interpretation.

Later we'll learn what makes calculus so fundamental in science and engineering. Derivatives are important in all measurements – in science, in engineering, in economics, in political science, in polling, in lots of commercial applications, in just about everything.

In this unit we'll also learn how to differentiate any function you know. That's a tall order, but by the end of the unit you will know how to take derivatives of functions like $f(x) = e^{x \cdot \arctan(x)}$.

Let's begin.

Geometric Interpretation of Differentiation

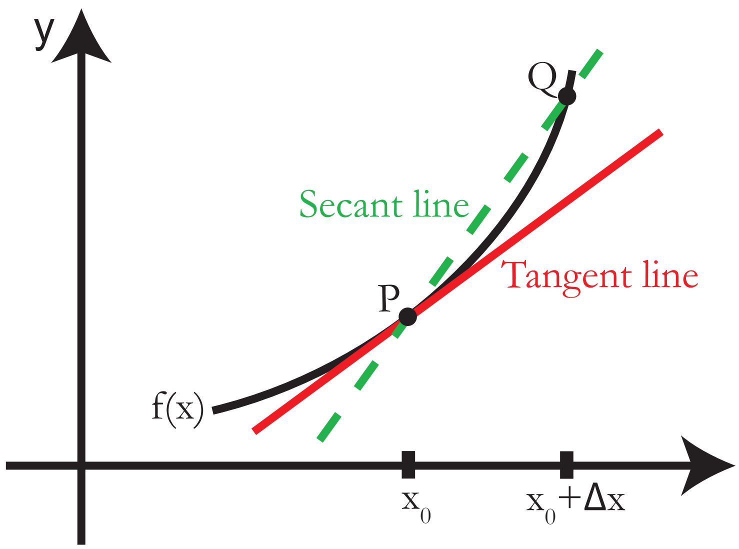

Figure 1: A graph with secant and tangent lines.

The derivative of $f(x)$ at $x = x_0$ is the slope of the tangent line to the graph of $f(x)$ at the point $(x_0, f(x_0))$. But what is a tangent line?

- It is NOT just a line that meets the graph at one point.
- It is the *limit* of the secant lines joining points $P = (x_0, f(x_0))$ and Q on the graph of $f(x)$ as Q approaches P.

The tangent line touches the graph at $(x_0, f(x_0))$; the slope of the tangent line matches the direction of the graph at that point. The tangent line is the straight line that best approximates the graph at that point.

Given a graph of our function, it's not hard for us to draw the tangent line to the graph. However, we'll want to do computations involving the tangent line and so will need a computational method of finding the tangent line.

How do we compute the equation of the line tangent to the graph of the function $f(x)$ at a point $P = (x_0, y_0)$? We know that the equation of the straight line with slope m through the point (x_0, y_0) is $y - y_0 = m(x - x_0)$, so in the abstract we know the equation of the tangent line.

To get a specific equation for the line, we'll need to know the coordinates x_0 and y_0 of the point P. If we know x_0 we can find $y_0 = f(x_0)$ by substituting the value x_0 in to the expression for $f(x)$. The second thing we need to know is the slope, $m = f'(x_0)$, which we call *the derivative of* f.

Definition: The derivative $f'(x_0)$ of f at x_0 is the slope of the tangent line to $y = f(x)$ at the point $P = (x_0, f(x_0)$.

Geometric definition of the derivative:

We're still trying to find a computational method of finding the equation of the tangent line – how do we compute the value of m?

In general, how do we know which lines are tangent lines and which lines are not?

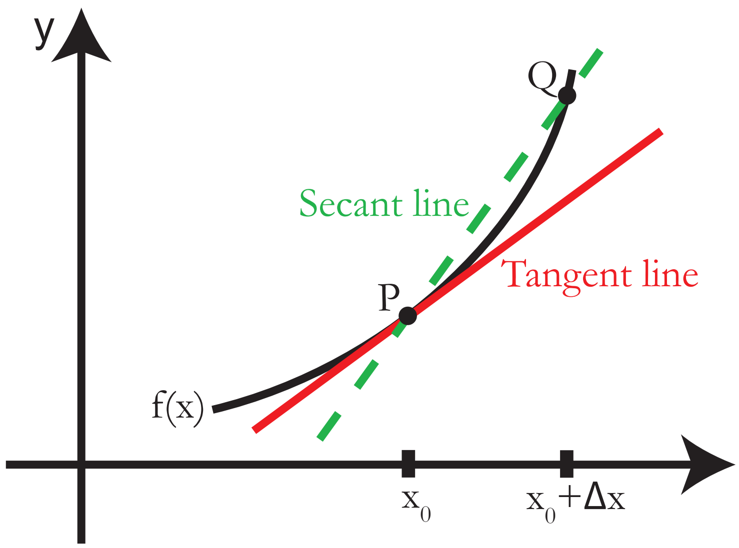

Figure 1: A graph with secant and tangent lines

A *secant* line is a line that joins two points on a curve. If the two points are close enough together, the slope of the secant line is close to the slope of the curve. We want to find the slope of the tangent line m — which equals the slope of the curve — and we use the slopes of secant lines to do this.

Suppose PQ is a secant line of the graph of $f(x)$. We can find the slope of the graph at P by calculating the slope of PQ as Q moves closer and closer to P (and the slope of PQ gets closer and closer to m).

The tangent line equals the limit of secant lines PQ as $Q \to P$; here P is fixed and Q varies.

Slope as Ratio

While we're still thinking geometrically, we can now use symbols and formulas in our computation.

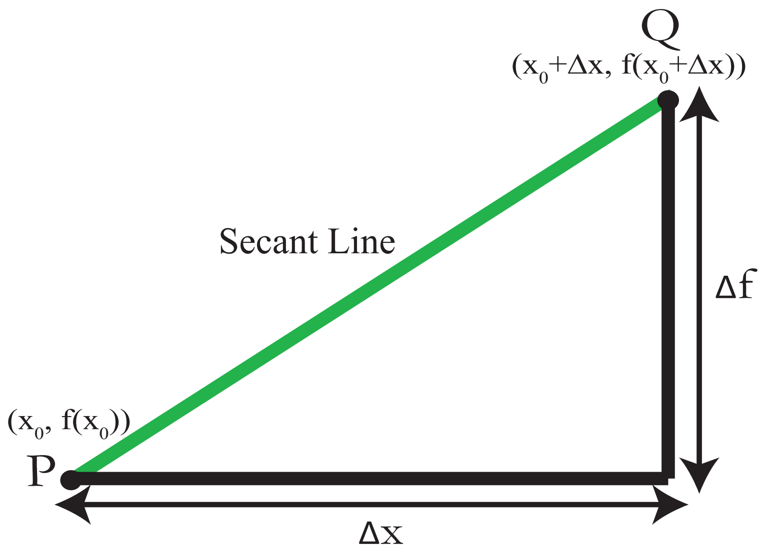

Figure 1: Geometric definition of the derivative

We start with a point $P = (x_0, f(x_0))$. We move over a tiny horizontal distance Δx (pronounced "delta x" and also called "the change in x") and find point $Q = (x_0 + \Delta x, f(x_0 + \Delta x))$. These two points lie on a secant line of the graph of $f(x)$; we will compute the slope of this line. The vertical difference between P and Q is $\Delta f = f(x_0 + \Delta x) - f(x_0)$.

The slope of the secant PQ is rise divided by run, or the ratio $\frac{\Delta f}{\Delta x}$. We've said that the tangent line is the limit of the secant lines. It is also true that the slope of the tangent line is the limit of the slopes of the secant lines. In other words,

$$m = \lim_{Q->P} \frac{\Delta f}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}.$$

Main Formula

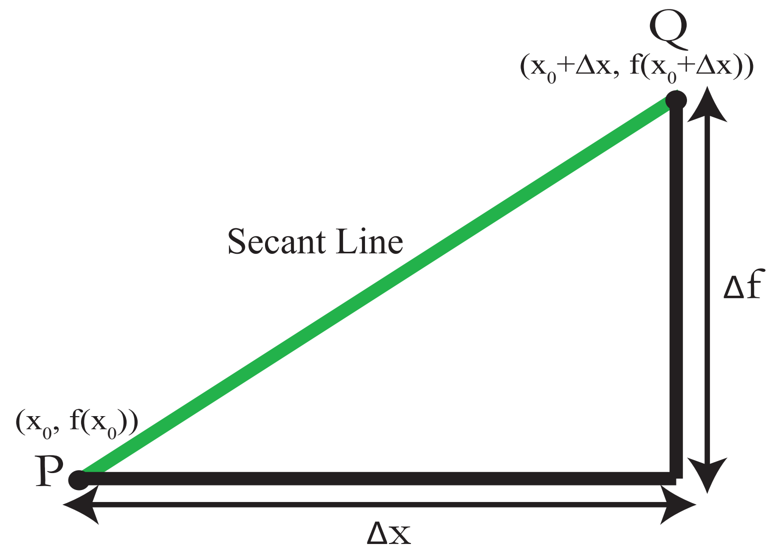

Figure 1: Geometric definition of the derivative

We started with a point P on the graph of $y = f(x)$ which had coordinates $(x_0, f(x_0))$. We then found a point Q on the the graph which was Δx units to the right of P. The coordinates of Q must be $(x_0 + \Delta x, f(x_0 + \Delta x))$. We can now write the following formula for the derivative:

$$m = \underbrace{f'(x_0)}_{\text{derivative of } f \text{ at } x_0} = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = \lim_{\Delta x \to 0} \underbrace{\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}}_{\text{difference quotient}}$$

This is by far the most important formula in Lecture 1; it is the formula that we use to compute the derivative $f'(x_0)$, which equals the slope of the tangent line to the graph at P. A machine could use this formula together with the coordinates $(x_0, f(x_0))$ of the point P to draw the tangent line to the graph of $y = f(x)$ at the point P.

Example 1. $f(x) = \dfrac{1}{x}$

We'll find the derivative of the function $f(x) = \frac{1}{x}$. To do this we will use the formula:

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Graphically, we will be finding the slope of the tangent line at at an arbitrary point $(x_0, \frac{1}{x_0})$ on the graph of $y = \frac{1}{x}$. (The graph of $y = \frac{1}{x}$ is a hyperbola in the same way that the graph of $y = x^2$ is a parabola.)

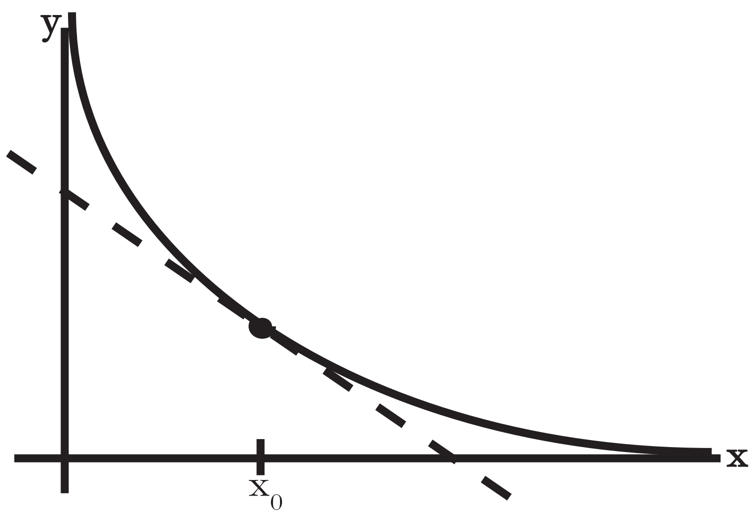

Figure 1: Graph of $\frac{1}{x}$

We start by computing the slope of the secant line:

$$\begin{aligned}
\frac{\Delta f}{\Delta x} &= \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \\
&= \frac{\frac{1}{x_0 + \Delta x} - \frac{1}{x_0}}{\Delta x} \\
&= \frac{(x_0)(x_0 + \Delta x)}{(x_0)(x_0 + \Delta x)} \frac{\frac{1}{x_0 + \Delta x} - \frac{1}{x_0}}{\Delta x} \\
&= \frac{\frac{(x_0)(x_0 + \Delta x)}{x_0 + \Delta x} - \frac{(x_0)(x_0 + \Delta x)}{x_0}}{(x_0)(x_0 + \Delta x)\Delta x} \\
&= \frac{1}{\Delta x} \frac{x_0 - (x_0 + \Delta x)}{(x_0)(x_0 + \Delta x)}
\end{aligned}$$

$$= \frac{1}{\Delta x}\frac{-\Delta x}{(x_0)(x_0+\Delta x)}$$
$$= \frac{-1}{(x_0)(x_0+\Delta x)}.$$

Next, we see what happens to the slopes of the secant lines as Δx tends to zero:

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = \lim_{\Delta x \to 0} \frac{-1}{(x_0)(x_0+\Delta x)} = \frac{-1}{x_0^2}$$

One thing to keep in mind when working with derivatives: it may be tempting to plug in $\Delta x = 0$ right away. If you do this, however, you will always end up with $\frac{\Delta f}{\Delta x} = \frac{0}{0}$. You will always need to do some cancellation to get at the answer.

We've computed that $f'(x) = \frac{-1}{x_0^2}$. Is this correct? How might we check our work? First of all, $f'(x_0)$ is negative — as is the slope of the tangent line on the graph of $y = \frac{1}{x}$. Secondly, as $x_0 \to \infty$ (i.e. as x_0 grows larger and larger), the tangent line is less and less steep. So $\frac{1}{x_0^2}$ should get closer to 0 as x_0 increases, which it does.

Question: Explain why $\lim_{\Delta x \to 0} \frac{-1}{(x_0)(x_0+\Delta x)} = \frac{-1}{x_0^2}$ again?

Answer: The point x_0 could be any point; let's suppose that $x_0 = 3$ so that we can look at this limit in a specific case.

We want to know the value of $\frac{-1}{(3)(3+\Delta x)}$ as Δx tends toward zero. As Δx gets smaller and smaller $3 + \Delta x$ gets closer and closer to 3, and so $\frac{-1}{(3)(3+\Delta x)}$ gets closer and closer to $\frac{-1}{(3)(3)} = \frac{-1}{9}$.

Question: Why is it that $\frac{\frac{1}{x_0+\Delta x} - \frac{1}{x_0}}{\Delta x} = \frac{1}{\Delta x}\frac{x_0 - (x_0+\Delta x)}{(x_0)(x_0+\Delta x)}$?

Answer: There are two steps in this simplification. We factored out the Δx that was in the denominator to become the $\frac{1}{\Delta x}$ "out front". At the same time, we rewrote the difference of two fractions $\frac{1}{x_0+\Delta x} - \frac{1}{x_0}$ using a common denominator.

This common denominator was $(x_0)(x_0 + \Delta x)$, which is just the product of the denominators in $\frac{1}{x_0+\Delta x} - \frac{1}{x_0}$. To get the common denominator, we multiply the first fraction by $\frac{(x_0)}{(x_0)} = 1$ and the second by $\frac{(x_0+\Delta x)}{(x_0+\Delta x)}$. (Multiplying by 1 won't change its value, but can change the algebraic expression we use to describe that value.) The denominators cancel, as intended, and we're left with $\frac{x_0-(x_0+\Delta x)}{(x_0)(x_0+\Delta x)}$.

A “Harder” Problem

People say that calculus is hard, but the example we just saw — computing the derivative of $f(x) = \frac{1}{x}$ — was not very difficult. What makes calculus seem hard is the context calculus problems appear in. For example, the problem we are about to solve combines algebra, geometry and problem solving with calculus. Because we use calculus to solve it, it is “a calculus problem”. And although it is a harder problem, it’s not the calculus that makes it hard.

So far all we’ve talked about is geometry, so our example problem must be geometric.

Problem: Find the area of the triangle formed by the x and y-axes and the tangent to the graph of $y = \frac{1}{x}$.

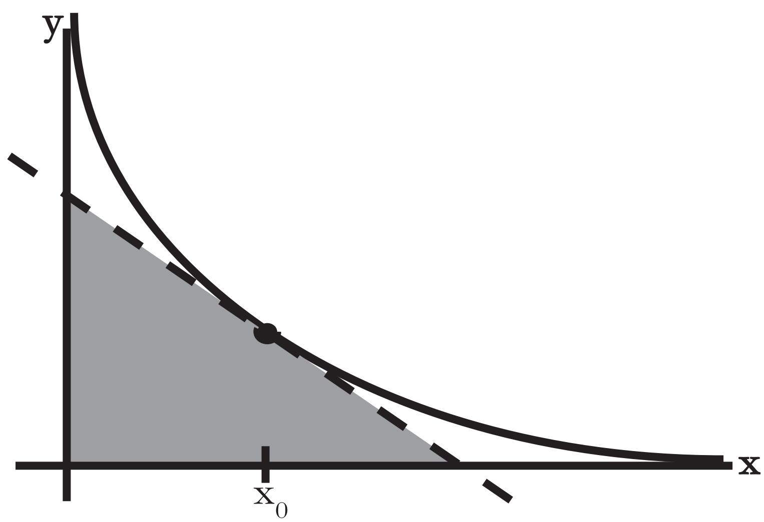

Figure 1: Triangle formed by axes and tangent line

We start by drawing a picture. As we draw the picture, we realize we’ve assumed that the triangle lies in the first quadrant. The solution we find might or might not depend on this assumption — we’ll have to do some work later if we want to be sure it’s true when x and $\frac{1}{x}$ are negative.

Because the problem refers to a tangent line it is a calculus problem, but as you’ll see, the calculus is the easy part.

The next step in our solution is labeling the picture. We label the graph $y = \frac{1}{x}$. We’d like to put some labels on the triangle, like the lengths of its sides, but we don’t know how to find those numbers. What we do know is that the hypotenuse of the triangle is tangent to the graph, and we can label the point of tangency $(x_0, \frac{1}{x_0})$ (which lets us label the points $(x_0, 0)$ and $(0, y_0)$ on the axes).

The area of a triangle is $\frac{1}{2}b \cdot h$, so we’d like to find the lengths of the base and height of this triangle. If we can find the x-coordinate of the point where the tangent line intersects the x-axis, we'll know the length of the base of the triangle. Similarly, finding the y-intercept of the tangent line will give us the triangle's height.

In order to find the x- and y-intercepts of the tangent line, we first must find the equation of the line. We start by writing down the "point-slope" form of the equation for a line:

$$y - y_0 = m(x - x_0).$$

The part of this problem that requires calculus is finding the slope m of the tangent line. Luckily, we already did this calculation in the previous example: $m = -\frac{1}{x_0^2}$.

$$y - y_0 = -\frac{1}{x_0^2}(x - x_0)$$

We've finished all the calculus in the problem, but we still need to do some work to find the area of the triangle. The point (x_0, y_0) is the point where the hypotenuse is tangent to the graph of $y = \frac{1}{x}$. We leave x_0 as is, and replace y_0 by $\frac{1}{x_0}$.

$$y - \frac{1}{x_0} = -\frac{1}{x_0^2}(x - x_0).$$

First we calculate the x-intercept of the tangent line, which will give us the length of the base of the triangle. The line crosses the x-axis when $y = 0$. Setting $y = 0$ in the equation for the tangent line, we get:

$$\begin{aligned}
0 - \frac{1}{x_0} &= \frac{-1}{x_0^2}(x - x_0) \\
\frac{-1}{x_0} &= \frac{-1}{x_0^2}x + \frac{1}{x_0} \\
\frac{1}{x_0^2}x &= \frac{2}{x_0} \\
x &= x_0^2(\frac{2}{x_0}) \;=\; 2x_0
\end{aligned}$$

So, the x-intercept of this tangent line is at $x = 2x_0$.

Next we could find the y-intercept using a very similar calculation — replacing x by 0 and solving for y (because the y-intercept is the point on the line with x coordinate 0). If we don't want to do the calculation all over again, or if we like using clever tricks to make problems easier, we can use the symmetry of the graph to take a short cut.

Since $y = \dfrac{1}{x}$ and $x = \dfrac{1}{y}$ are identical equations, the graph of $y = \frac{1}{x}$ is symmetric when x and y are exchanged. By symmetry, then, we could swap the

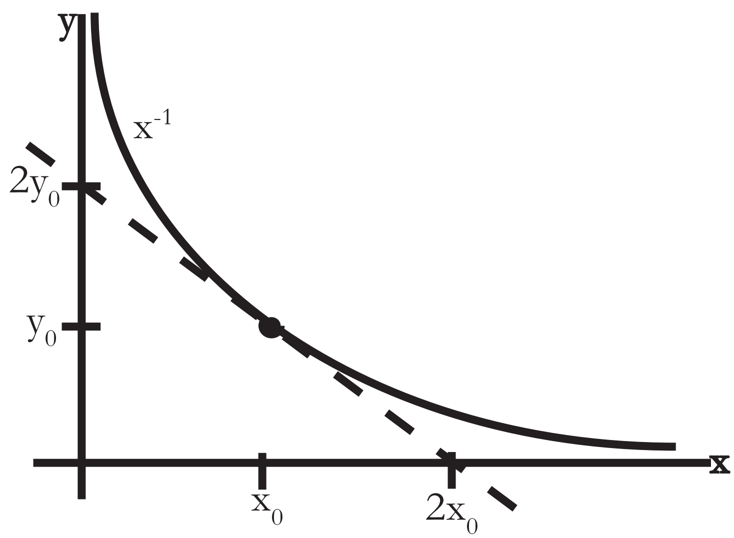

Figure 2: Triangle formed by axes and tangent line, labeled

x's in the calculation above with the y's to conclude that the y-intercept is at $y = 2y_0 = \frac{2}{x_0}$.

Finally,

$$\text{Area} = \frac{1}{2}\text{b} \cdot \text{h} = \frac{1}{2}(2\text{x}_0)(2\text{y}_0) = 2\text{x}_0\text{y}_0 = 2\text{x}_0(\frac{1}{\text{x}_0}) = 2 \text{ (see Fig. 2)}$$

Curiously, the area of the triangle is *always* 2, no matter where on the graph we draw the tangent line!

Remark: We call it "one variable calculus", but we just used *four* variables: x, y, x_0, and y_0. We could have had more! This makes things complicated, and it's something that you'll have to get used to.

Another complicated thing that we do is reuse variables. In this problem, there are (at least) three different possible interpretations of the variable y. When we said $y = \frac{1}{x}$, we were thinking of y as the vertical position of a point on the hyperbola. When we said $y - y_0 = \frac{-1}{x_0^2}(x - x_0)$ we were thinking of y as the vertical position of a point on the tangent line. And when we said $y = 0$ we were talking about the vertical position of all the points on the x-axis. Once you've practiced calculus for a while you will know from context which meaning y has, just as you can tell in conversation whether a person is saying "sea" or "see". Until you reach that point, be sure you understand the meaning of an equation before using it in a calculation.

Notations

Calculus, rather like English or any other language, was developed by several people. As a result, just as there are many ways to express the same thing, there are many notations for the derivative.

Since $y = f(x)$, it's natural to write

$$\Delta y = \Delta f = f(x) - f(x_0) = f(x_0 + \Delta x) - f(x_0)$$

We say "Delta y" or "Delta f" or the "change in y".

If we divide both sides by $\Delta x = x - x_0$, we get two expressions for the difference quotient:

$$\frac{\Delta y}{\Delta x} = \frac{\Delta f}{\Delta x}$$

Taking the limit as $\Delta x \to 0$, we get

$$\begin{aligned} \frac{\Delta y}{\Delta x} \quad &\to \quad \frac{dy}{dx} \text{ (Leibniz' notation)} \\ \frac{\Delta f}{\Delta x} \quad &\to \quad f'(x_0) \text{ (Newton's notation)} \end{aligned}$$

In Leibniz' notation we might also write $\frac{df}{dx}$, $\frac{d}{dx}f$ or $\frac{d}{dx}y$. Notice that Leibniz' notation doesn't specify where you're evaluating the derivative. In the example of $f(x) = \frac{1}{x}$ we were evaluating the derivative at $x = x_0$.

Other, equally valid notations for the derivative of a function f include f' and Df.

Example 2. $f(x) = x^n$ where $n = 1, 2, 3...$

In this example we answer the question "What is $\frac{d}{dx}x^n$?" Once we know the answer we can use it to, for example, find the derivative of $f(x) = x^4$ by replacing n by 4.

At this point in our studies, we only know one tool for finding derivatives – the difference quotient. So we plug $y = f(x)$ into the definition of the difference quotient:

$$\frac{\Delta y}{\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \frac{(x_0 + \Delta x)^n - x_0^n}{\Delta x}$$

Because writing little zeros under all our x's is a nuisance and a waste of chalk (or of photons?), and because there's no other variable named x to get confused with, from here on we'll replace x_0 with x.

$$\frac{\Delta y}{\Delta x} = \frac{(x + \Delta x)^n - x^n}{\Delta x}$$

Remember that when we use the difference quotient, we're thinking of x as fixed and of Δx as getting closer to zero. We want to simplify this fraction so that we can plug in 0 for Δx without any danger of dividing by zero. To do this we must expand the expression $(x + \Delta x)^n$.

A famous formula called the binomial theorem tells us that:

$$(x + \Delta x)^n = (x + \Delta x)(x + \Delta x)...(x + \Delta x) \quad \text{n times}$$

We can rewrite this as

$$x^n + n(\Delta x)x^{n-1} + O\left((\Delta x)^2\right)$$

where $O(\Delta x)^2$ is shorthand for "all of the terms with $(\Delta x)^2$, $(\Delta x)^3$, and so on up to $(\Delta x)^n$."

One way to begin to understand this is to think about multiplying all the x's together from

$$(x + \Delta x)^n = (x + \Delta x)(x + \Delta x)...(x + \Delta x) \quad \text{n times.}$$

There are n of these x's, so multiplying them together gives you one term of x^n. What if you only multiply together $n - 1$ of the x's? Then you have one $(x+\Delta x)$ left that you haven't taken an x from, and you can multiply your x^{n-1} by Δx. (If you multiplied by x, you'd just have the x^n that you already got.) There were n different Δx's that you could have chosen to use, so you can get this result n different ways. That's where the $n(\Delta x)x^{n-1}$ comes from.

We could keep going, and figure out how many different ways there are to multiply $n - 2$ x's by two Δx's, and so on, but it turns out we don't need to. Every other way of multiplying together one thing from each $(x + \Delta x)$ gives you at least two Δx's, and $\Delta x \cdot \Delta x$ is going to be too small to matter to us as $\Delta x \to 0$.

Now that we have some idea of what $(x + \Delta x)^n$ is, let's go back to our difference quotient.

$$\frac{\Delta y}{\Delta x} = \frac{(x + \Delta x)^n - x^n}{\Delta x} = \frac{(x^n + n(\Delta x)(x^{n-1}) + O(\Delta x)^2) - x^n}{\Delta x} = nx^{n-1} + O(\Delta x)$$

As it turns out, we *can* simplify the quotient by canceling a Δx in all of the terms in the numerator. When we divide a term that contains Δx^2 by Δx, the Δx^2 becomes Δx and so our $O(\Delta x^2)$ becomes $O(\Delta x)$.

When we take the limit as x approaches 0 we get:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = nx^{n-1}$$

and therefore,

$$\frac{d}{dx} x^n = nx^{n-1}$$

This result is sometimes called the "power rule". We will use it often to find derivatives of polynomials; for example,

$$\frac{d}{dx}(x^2 + 3x^{10}) = 2x + 30x^9$$

Introduction to Rates of Change

Last class, we defined the derivative as the slope of a tangent line. Today we'll see how to interpret the derivative as a rate of change, clarify the idea of a limit, and use this notion of limit to describe continuity – a property functions need to have in order for us to work with them.

Rates of Change

Last class we talked about the derivative as the slope of the tangent line to a graph. This class we'll continue our discussion of derivatives by explaining how a derivative can be a rate of change. This some of the most important information presented in this class.

Remember that when we talked about the slope of a graph $y = f(x)$ we started by talking about the change in y and the change in x. If changing x at a certain rate causes y to change, we're interested in the *relative* rate of change, $\frac{\Delta y}{\Delta x}$.

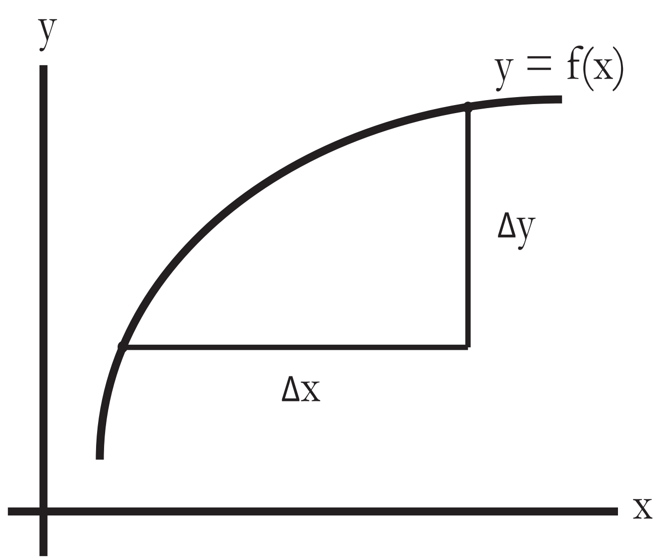

Figure 1: Graph of a generic function, with Δx and Δy marked on the graph

Another way to think about $\frac{\Delta y}{\Delta x}$ is as the average change in y over an interval of size Δx. This comes up frequently in physics, in which x is measuring time and $\frac{\Delta y}{\Delta x}$ is the average change in position over an interval of time – in other words, it's the rate at which something is moving. In this case, the limit

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$$

measures the instantaneous rate of change, or the speed.

Physical Interpretation of Derivatives

You can think of the derivative as representing a rate of change (speed is one example of this). This makes it very useful for solving physics problems.

Here's one example from physics: If q is an amount of electric charge, the derivative $\frac{dq}{dt}$ is the change in that charge over time, or the electric current.

A second, more tangible example is to let s stand for distance; then the rate of change $\frac{ds}{dt}$ is what we call speed. Let's investigate this second example in more detail to get a visceral sense of what instantaneous speed means.

On Halloween, MIT students have a tradition of dropping pumpkins from the roof of the building this lecture was given in. Let's say that the building is about 300 feet tall. We'll use a slightly smaller value of 80 meters for the height because it makes the problem easier to solve.

The equation of motion for objects near the earth's surface (which we will just accept for now) says that the height above the ground h of the pumpkin t seconds after it's dropped from the building is roughly:

$$h = 80 - 5t^2 \text{ meters}$$

Let's think about this. The instant the pumpkin is dropped, $t = 0$ and $h = 80$ meters. When $t = 4$ seconds, $h = 80 - 5(4^2) = 0$, and the pumpkin has reached the ground.

The average speed of the pumpkin over the time it's falling is

$$\frac{\Delta h}{\Delta t} = \frac{\text{distance traveled}}{\text{time elapsed}} = \frac{0 - 80 \text{ meters}}{4 - 0 \text{ seconds}} = -20 \text{ m/s}.$$

(The numerator is $0 - 80$ and not $80 - 0$ because we must subtract the initial position from the final position, not the other way around.)

The people watching the pumpkin drop probably don't care about the *average* speed. They want to know how fast the pumpkin is going when it slams into the ground. That's known as the instantaneous speed, and is the derivative $h'(t) = \frac{d}{dt}h$. To find the instantaneous velocity at $t = 5$, we evaluate $\frac{d}{dt}h$.

$$\frac{d}{dt}h = 0 - 10t = -10t$$

If you've had calculus before, you're probably able to find the derivative of the polynomial $80 - 5t^2$ on your own. If not, you'll have to take a few things on faith here. First, the derivative of $80 - 5t^2$ is just the derivative of 80 minus the derivative of $5t^2$. Next, the derivative of 80 is the slope of the graph of $y = 80$ when $x = 0$; that graph is a horizontal line! And finally, since we know from last class that the derivative of t^2 is $2t^{2-1} = 2t$ it should not surprise you that the derivative of $5t^2$ is $10t$.

We know that the pumpkin hits the pavement 4 seconds after it's dropped, at time $t = 4$, so the pumpkin's speed is:

$$h'(4) = (-10)(4) = -40 \text{ m/s} \quad \text{(about 90 mph or 145 kph)}.$$

The value of $\frac{d}{dt}h$ is negative because the pumpkin's height is decreasing; it is moving downward.

In actuality, the building is a little taller than 80 meters and there is air resistance. You may do a much more thorough study on your own if you wish.

Physical Interpretation of Derivatives, Continued

Calculus is a good tool for studying how things change over time; we can also use it to investigate change with respect to variables other than time. Let's look at a couple of examples that don't involve time as a variable.

If we let T denote temperature and x measure a distance or position, then $\frac{dT}{dx}$ is the rate the temperature changes as position changes. This quantity is called the *temperature gradient.* Temperature gradients are important in weather forecasting because it's that temperature difference that causes air flows and weather changes.

$T =$ temperature $\dfrac{dT}{dx} =$ temperature gradient

Another application of calculus is to "sensitivity of measurements". In Problem Set 1 there's a question about GPS and a satellite which explores sensitivity of measurements.

In the problem set, you assume that the earth is flat and that you have a satellite above a known location. A traveler's GPS device measures the distance h between the traveler and the satellite and uses that information to compute the horizontal distance L between the traveler and the point directly under the satellite. (See Fig. 1 and Fig. 2)

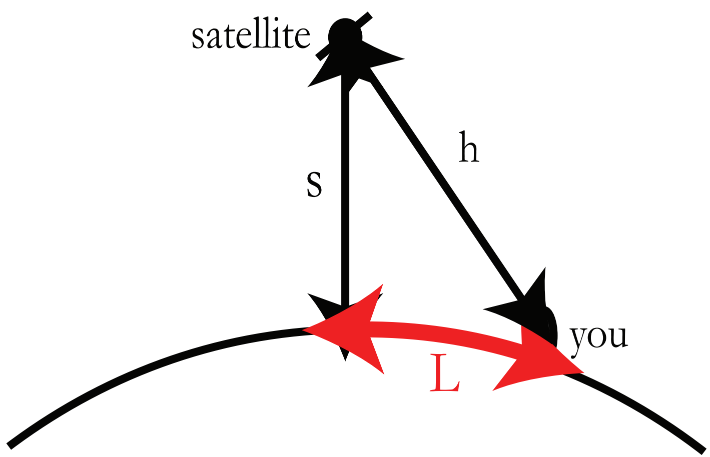

Figure 1: The Global Positioning System Problem (GPS)

In other words, the GPS computes L as a function of h. But there is usually some error Δh in your measurement of h; given that, how accurately can we measure L? The error ΔL is estimated by looking at $\dfrac{\Delta L}{\Delta h} \approx \dfrac{dL}{dh}$.

Why is this important? For one thing, it's used all the time to land airplanes.

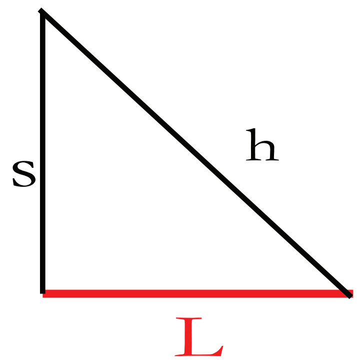

Figure 2: On problem set 1, you will look at this simplified “flat earth” model

This concludes our introduction to differentiation, although there will be plenty of opportunities throughout the course for you to improve your understanding.

Limits

Last class we talked about a series of secant lines approaching the "limit" of a tangent line, and about how as Δx approaches zero, $\frac{\Delta y}{\Delta x}$ approaches the "limit" $y' = \frac{dy}{dx}$. Now we want to talk about limits more carefully; this will include some of our first steps towards our goal of being able to differentiate every function you know.

Some limits are easy to compute:

$$\lim_{x\to 3} \frac{x^2+x}{x+1} = \frac{3^2+3}{3+1} = \frac{12}{4} = 3$$

With an easy limit, you can get a meaningful answer just by plugging in the limiting value. This is because when x is close to 3, the value of the function $f(x) = \frac{x^2+x}{x+1}$ is close to $f(3)$.

Some limits are not easy to compute. For example, the definition of the derivative:

$$\lim_{x\to x_0} \frac{\Delta f}{\Delta x} = \lim_{x\to x_0} \frac{f(x_0+\Delta x) - f(x_0)}{\Delta x}$$

is never an easy limit, because the denominator $\Delta x = 0$ is not allowed. (The limit $x \to x_0$ is computed under the implicit assumption that $x \neq x_0$.) We'll always need to cancel Δx before we can make sense out of the limit.

Other "hard" limits would be:

$$\lim_{x\to -1} \frac{x^2+x}{x+1} \quad \text{and} \quad \lim_{x\to\infty} \frac{x^2+x}{x+1}.$$

Any limit involving infinity or division by zero is going to be harder to compute; sometimes the answer will will be that there is no limit.

To complete our discussion of limits, we need just one more piece of notation — the concepts of left hand and right hand limits.

The limit

$$\lim_{x\to x_0^+} f(x)$$

is known as the *right-hand limit* and means that you should use values of x that are greater than x_0 (to the right of x_0 on the number line) to compute the limit. Shown below is the graph of the function:

$$f(x) = \begin{cases} x+1 & x > 0 \\ -x & x \leq 0 \end{cases}$$

The right-hand limit $\lim_{x\to 0^+} f(x)$ equals 1.

The *left-hand limit*

$$\lim_{x\to x_0^-} f(x)$$

is found by looking at values of $f(x)$ when x is less than x_0 (to the left of x_0 on the number line). For this function, $\lim_{x\to 0^-} f(x) = 0$.

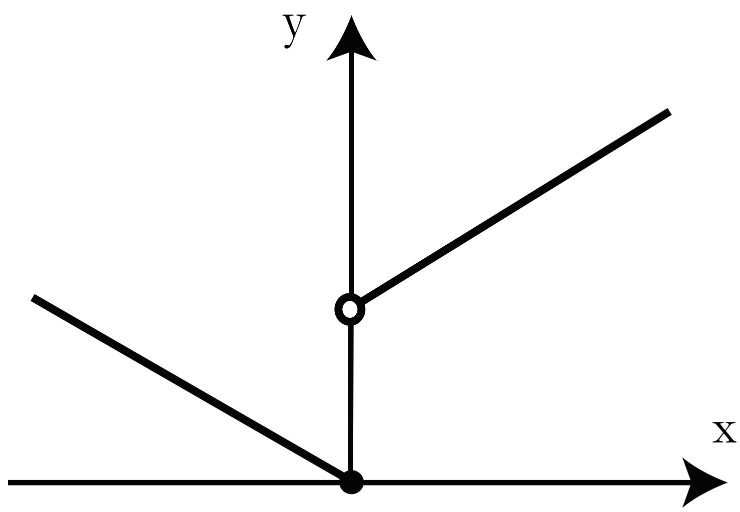

Figure 1: Graph of $f(x)$

The notions of left- and right- hand limits will make things much easier for us as we discuss continuity, next.

Let's talk more about the example graphed above. To calculate

$$\lim_{x \to x_0^+} f(x)$$

we use only values of x that are greater than 0. When $x > 0$, $f(x)$ is defined to equal $x + 1$. So we plugged $x = 0$ into the expression $x + 1$ to calculate the right-hand limit.

When calculating

$$\lim_{x \to x_0^-} f(x),$$

we have $x < 0$. Here $f(x)$ is defined to equal $-x$; when we plug $x = 0$ into this expression we get $\lim_{x \to x_0^-} f(x) = 0$.

Notice that it doesn't matter that $f(0) = 0$. Our calculations would have been exactly the same if $f(0)$ were 1 or even if $f(0) = 2$.

Continuity

Continuous Functions

Definition: A function f is *continuous* at x_0 if $\lim_{x \to x_0} f(x) = f(x_0)$.

What is this definition saying? A function that's continuous at x_0 has the following properties:

- $\lim_{x \to x_0^+} f(x) = \lim_{x \to x_0^-} f(x)$; in particular, both of these one sided limits exist.
- $f(x_0)$ is defined.
- $\lim_{x \to x_0^+} f(x) = \lim_{x \to x_0^-} f(x) = f(x_0)$.

This may look obvious, but remember that when you are calculating $\lim\limits_{x \to x_0} f(x)$ you never allow x to equal x_0. The value $\lim\limits_{x \to x_0} f(x)$ is computed independently of, and in a different way than, the value of $f(x_0)$. If we aren't careful to make this distinction, this definition has no meaning.

The limits for which $\lim\limits_{x \to x_0} f(x) = f(x_0)$ are exactly the "easy limits" we discussed earlier. The "harder" limits only happen for functions that are not continuous.

Next we'll see a tour of different types of discontinuous functions. The question of whether something is continuous or not may seem fussy, but it is something people have worried about a lot. Bob Merton, who was a professor at MIT when he did his work for the Nobel Prize in Economics, was interested in whether stock prices of various kinds are continuous from the left (past) or right (future) in a certain model. That was a serious consideration when developing a model that hedge funds now use all the time.

Jump Discontinuity

A *jump* discontinuity occurs when the right-hand and left-hand limits exist but are not equal. We've already seen one example of a function with a jump discontinuity:

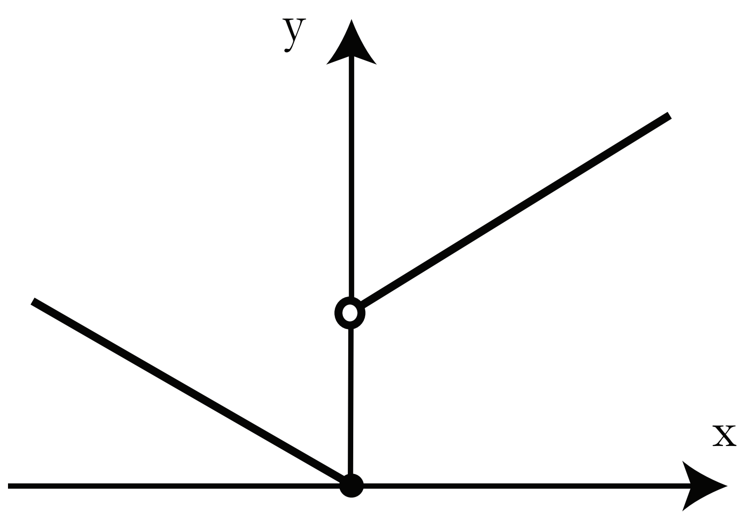

Figure 1: Graph of the discontinuous function listed below

$$f(x) = \begin{cases} x+1 & x > 0 \\ -x & x \geq 0 \end{cases}$$

This *discontinuous* function is seen in Fig. 1. For $x > 0$,

$$\lim_{x \to 0} f(x) = 1$$

but $f(0) = 0$. (One can also say, f is continuous from the left at 0, but not the right.)

Here is another example in which $\lim_{x \to x_0^+}$ exists, and $\lim_{x \to x_0^-}$ also exists, but they are NOT equal.

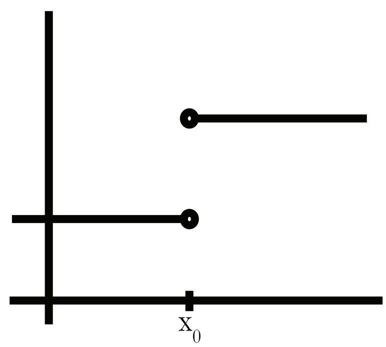

Figure 2: Another example of a jump discontinuity

Removable Discontinuities

At a *removable* discontinuity, the left-hand and right-hand limits are equal but either the function is not defined or not equal to these limits:

$$\lim_{x \to x_0^+} f(x) = \lim_{x \to x_0^-} f(x) \neq f(x_0)$$

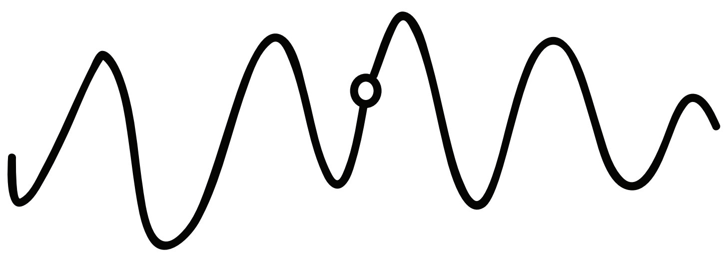

Figure 1: A removable discontinuity: the function is continuous everywhere except one point

For example, $g(x) = \dfrac{\sin(x)}{x}$ and $h(x) = \frac{1-\cos x}{x}$ are defined for $x \neq 0$, but both functions have removable discontinuities. This is not obvious at all, but we will learn later that:

$$\lim_{x \to 0} \frac{\sin x}{x} = 1 \text{ and } \lim_{x \to 0} \frac{1 - \cos x}{x} = 0.$$

So both of these functions have removable discontinuities at $x = 0$ despite the fact that the fractions defining them have a denominator of 0 when $x = 0$.

Infinite Discontinuities

In an *infinite* discontinuity, the left- and right-hand limits are infinite; they may be both positive, both negative, or one positive and one negative.

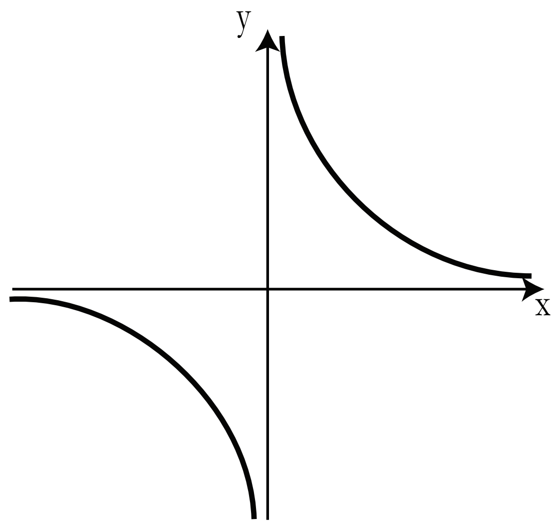

Figure 1: An example of an infinite discontinuity: $\dfrac{1}{x}$

From Figure 1, we see that $\lim\limits_{x\to 0^+} \dfrac{1}{x} = \infty$ and $\lim\limits_{x\to 0^-} \dfrac{1}{x} = -\infty$. Saying that a limit is ∞ is different from saying that the limit doesn't exist – the values of $\frac{1}{x}$ are changing in a very definite way as $x \to 0$ from either side. (Note that it's not true that $\lim\limits_{x\to 0} \dfrac{1}{x} = \infty$ because ∞ and $-\infty$ are different.)

There are two more things we can learn from this example. First, sketch the graph of $\frac{d}{dx}\frac{1}{x} = -\frac{1}{x^2}$; it also has an infinite discontinuity at $x = 0$. Notice that the derivative of the function $\frac{1}{x}$ is always negative. It may seem strange to you that the derivative is decreasing as x approaches 0 from the positive side while $\frac{1}{x}$ is increasing, but very often the graph of the derivative will look nothing like the graph of the original function.

What the graph of the derivative $-\frac{1}{x^2}$ is showing you is the slope of the graph of $\frac{1}{x}$. Where the graph of $\frac{1}{x}$ is not very steep, the graph of $-\frac{1}{x^2}$ lies close to the x-axis. Where the graph of $\frac{1}{x}$ is steep, the graph of $-\frac{1}{x^2}$ is far away from the x-axis. The value of $-\frac{1}{x^2}$ is always negative, and the graph of $\frac{1}{x}$ always slopes downward.

Finally, $\frac{1}{x}$ is an odd function and $-\frac{1}{x^2}$ is an even function. When you take the derivative of an odd function you always get an even function and vice-versa. If you can easily identify odd and even functions, this is a good way to check

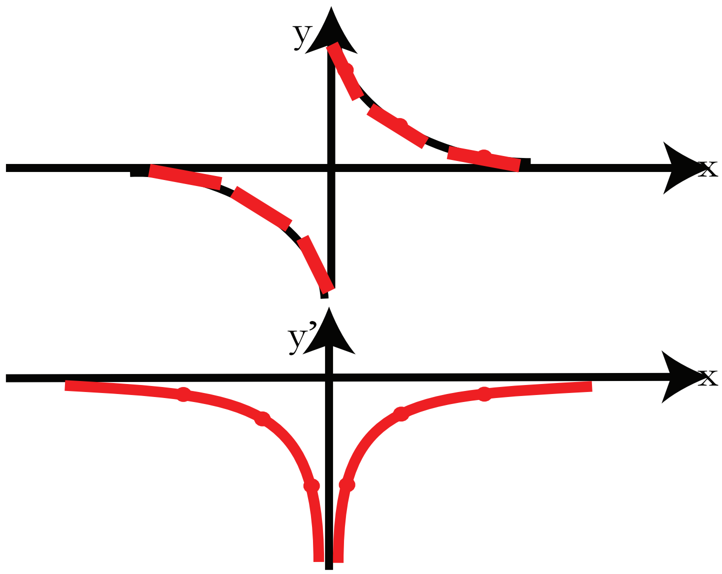

Figure 2: Top: graph of $f(x) = \dfrac{1}{x}$ and Bottom: graph of $f'(x) = -\dfrac{1}{x^2}$

your work.

Other (Ugly) Discontinuities

The limit $\lim\limits_{x \to 0} \sin(1/x)$ *is* undefined as x goes to 0. The graph of $y = \sin(1/x)$ is similar to the one in Figure 1; the function $\sin(1/x)$ has no left or right limit as x goes to 0. Here, we say the limit does not exist.

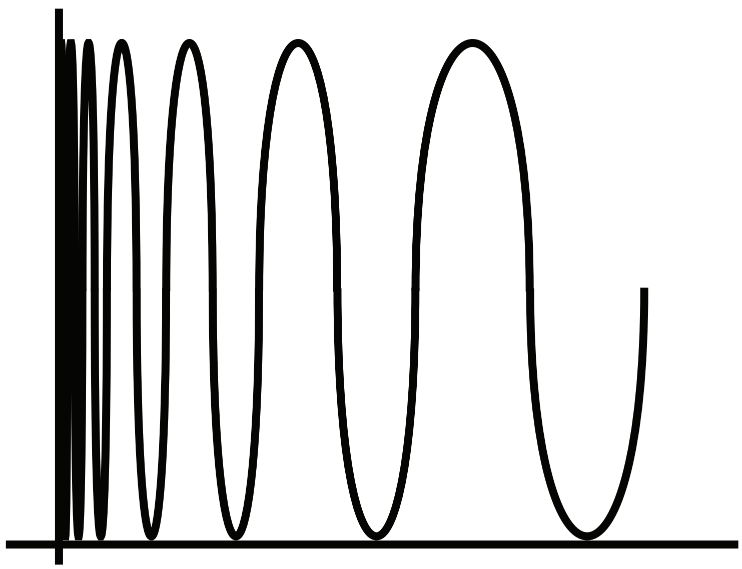

Figure 1: An example of an ugly discontinuity: a function that oscillates a lot as it approaches the origin

There are many discontinuities of this type — for example, things that oscillate as time goes to infinity — but we're not going to worry about them in this course.

Differentiable Implies Continuous

Theorem: If f is differentiable at x_0, then f is continuous at x_0.

We need to prove this theorem so that we can use it to find general formulas for products and quotients of functions.

We begin by writing down what we need to prove; we choose this carefully to make the rest of the proof easier. We want to show that:

$$\lim_{x \to x_0} f(x) - f(x_0) = 0.$$

This is the same as saying that the function is continuous, because to prove that a function was continuous we'd show that $\lim_{x \to x_0} f(x) = f(x_0)$.

We prove $\lim_{x \to x_0} f(x) - f(x_0) = 0$ by multiplying and dividing it by the same number – this won't change its value.

$$\begin{aligned} \lim_{x \to x_0} f(x) - f(x_0) &= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}(x - x_0) \\ &= f'(x) \cdot 0 \\ &= 0. \end{aligned}$$

(Notice that we used our assumption that f was differentiable when we wrote down $f'(x)$.)

But wait! When we multiplied and divided by $x - x_0$ weren't we multiplying and dividing by zero? We know from our algebra classes that this never works! It turns out that we're safe because we're using limits. Although x gets closer and closer to x_0, it never actually equals x_0, and so we never quite divide by 0. That's what limits are for; $x - x_0$ may be small, but it's always non-zero.

So this calculation is valid, it's true that $\lim_{x \to x_0} f(x) - f(x_0) = 0$, and it's true that differentiable functions are continuous.

Introduction to Differentiation

Working toward our goal of "differentiating everything", this lecture introduces some useful new formulas.

There are two basic types of derivative formulas:

1. Specific Examples: power rule

2. General Examples: (u +v)' = u'+v' and (cu)' = cu' (where c is a constant)

We need both kinds of formulas to take derivatives of polynomials, for example.

This lecture focuses on the basic trig functions, finding specific formulas for the derivative of the sine function and the cosine function.

Derivative of a Sum

One of our examples of a general derivative formula was:

$$(u+v)'(x) = u'(x) + v'(x).$$

(Remember that by $(u+v)(x)$ we mean $u(x)+v(x)$.)

In other words, the derivative of the sum of two functions is just the sum of their derivatives. We'll now prove that this is true for any pair of functions u and v, provided that those functions have derivatives. Since we don't know in advance what functions u and v are, we can't use any specific information about the functions or the slopes of their graphs; all we have to work with is the formal definition of the derivative.

When we apply the definition of the derivative to the function $(u+v)(x)$ we get:

$$(u+v)'(x) = \lim_{\Delta x \to 0} \frac{(u+v)(x+\Delta x) - (u+v)(x)}{\Delta x}$$

Since $(u+v)(x)$ is just $u(x)+v(x)$,

$$(u+v)'(x) = \lim_{\Delta x \to 0} \frac{u(x+\Delta x) + v(x+\Delta x) - u(x) - v(x)}{\Delta x}.$$

Combining like terms, we see that:

$$(u+v)'(x) = \lim_{\Delta x \to 0} \frac{u(x+\Delta x) - u(x) + v(x+\Delta x) - v(x)}{\Delta x}$$

or:

$$(u+v)'(x) = \lim_{\Delta x \to 0} \left\{ \frac{u(x+\Delta x) - u(x)}{\Delta x} + \frac{v(x+\Delta x) - v(x)}{\Delta x} \right\}.$$

Because u and v are differentiable (and therefore continuous), the limit of the sum is the sum of the limits. Therefore:

$$(u+v)'(x) = \lim_{\Delta x \to 0} \frac{u(x+\Delta x) - u(x)}{\Delta x} + \lim_{\Delta x \to 0} \frac{v(x+\Delta x) - v(x)}{\Delta x}.$$

The two limits above match the definition of the derivatives of u and v, so we've shown that $(u+v)'(x) = u'(x) + v'(x)$.

Derivative of $\sin x$, Algebraic Proof

A specific derivative formula tells us how to take the derivative of a specific function: if $f(x) = x^n$ then $f'(x) = nx^{n-1}$. We'll now compute a specific formula for the derivative of the function $\sin x$.

As before, we begin with the definition of the derivative:

$$\frac{d}{dx}\sin x = \lim_{\Delta x \to 0} \frac{\sin(x + \Delta x) - \sin(x)}{\Delta x}$$

You may remember the following angle sum formula from high school:

$$\sin(a + b) = \sin(a)\cos(b) + \sin(b)\cos(a)$$

This lets us untangle the x from the Δx as follows:

$$\frac{d}{dx}\sin x = \lim_{\Delta x \to 0} \frac{\sin x \cos \Delta x + \cos x \sin \Delta x - \sin(x)}{\Delta x}.$$

We can simplify this expression using some basic algebraic facts:

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

We now have two familiar functions – $\sin x$ and $\cos x$ – and two ugly looking fractions to deal with. The fractions may be familiar from our discussion of removable discontinuities.

$$\begin{aligned}
\lim_{\Delta x \to 0} \frac{\cos \Delta x - 1}{\Delta x} &= 0 \\
\lim_{\Delta x \to 0} \frac{\sin \Delta x}{\Delta x} &= 1.
\end{aligned}$$

Using these (as yet unproven) facts,

$$\lim_{\Delta x \to 0} \sin x \left(\frac{\cos \Delta x - 1}{\Delta x}\right) + \lim_{\Delta x \to 0} \cos x \left(\frac{\sin \Delta x}{\Delta x}\right)$$

simplifies to $\sin x \cdot 0 + \cos x \cdot 1 = \cos x$. We conclude:

$$\frac{d}{dx}\sin x = \cos x$$

Derivative of $\cos x$.

What is the specific formula for the derivative of the function $\cos x$?

This calculation is very similar to that of the derivative of $\sin(x)$. If you get stuck on a step here it may help to go back and review the corresponding step there.

As in the calculation of $\frac{d}{dx}\sin x$, we begin with the definition of the derivative:

$$\frac{d}{dx}\cos x = \lim_{\Delta x \to 0} \frac{\cos(x + \Delta x) - \cos(x)}{\Delta x}$$

Use the angle sum formula $\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)$ and then simplify:

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

Once again we use the following (unproven) facts:

$$\begin{aligned}
\lim_{\Delta x \to 0} \frac{\cos \Delta x - 1}{\Delta x} &= 0 \text{ (A)} \\
\lim_{\Delta x \to 0} \frac{\sin \Delta x}{\Delta x} &= 1. \text{ (B)}
\end{aligned}$$

And we conclude:

$$\begin{aligned}
\frac{d}{dx}\cos x &= \lim_{\Delta x \to 0} \cos x \left(\frac{\cos \Delta x - 1}{\Delta x} \right) + \lim_{\Delta x \to 0} (-\sin x) \left(\frac{\sin \Delta x}{\Delta x} \right) \\
&= \cos x \cdot 0 + (-\sin x) \cdot 1 \\
\frac{d}{dx}\cos x &= -sin(x).
\end{aligned}$$

$$\lim_{x\to 0}\frac{sin(x)}{x} = 1$$

In order to compute specific formulas for the derivatives of $\sin(x)$ and $\cos(x)$, we needed to understand the behavior of $\sin(x)/x$ near $x = 0$ (property B). In his lecture, Professor Jerison uses the definition of $\sin(\theta)$ as the y-coordinate of a point on the unit circle to prove that $\lim_{\theta\to 0}(\sin(\theta)/\theta) = 1$.

We switch from using x to using θ because we want to start thinking about the sine function as describing a ratio of sides in the triangle shown in Figure 1. The variable we're interested in is an angle, not a horizontal position, so we discuss $\sin(\theta)/\theta$ rather than $\sin(x)/x$.

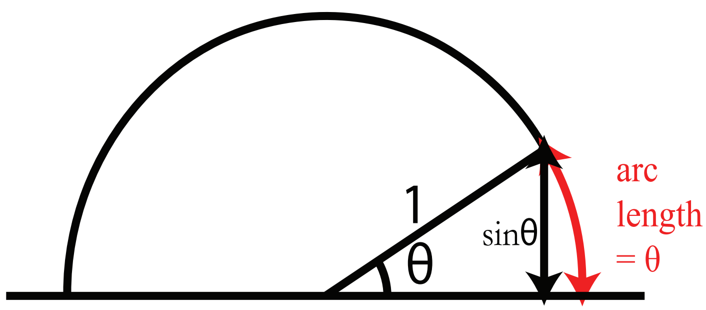

Figure 1: A circle of radius 1 with an arc of angle θ.

Our argument depends on the fact that when the radius of the circle shown in Figure 1 is 1, θ is the length of the highlighted arc. This is true when the angle θ is described in radians but NOT when it is measured in degrees.

Also, since the radius of the circle is 1, $\sin(\theta) = \frac{|\text{opposite}|}{|\text{hypotenuse}|}$ equals the length of the edge indicated (the hypotenuse has length 1).

In other words, $\sin(\theta)/\theta$ is the ratio of edge length to arc length. When $\theta = \pi/2$ rad, $\sin(\theta) = 1$ and $\sin(\theta)/\theta = 2/\pi \cong 2/3$. When $\theta = \pi/4$ rad, $\sin(\theta) = \sqrt{2}/2$ and $\sin(\theta)/\theta = 2\sqrt{2}/\pi \cong 9/10$. What will happen to the value of $\sin(\theta)/\theta$ as the value of θ gets closer and closer to 0 radians?

We see from Figure 2 that as θ shrinks, the length $\sin(\theta)$ of the segment gets closer and closer to the length θ of the curved arc. We conclude that as $\theta \to 0$,

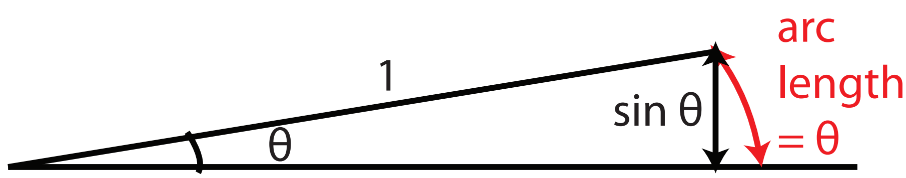

Figure 2: The sector in Fig. 1 as θ becomes very small

$\dfrac{\sin\theta}{\theta} \to 1$. In other words,

$$\lim_{x\to 0} \frac{\sin(x)}{x} = 1.$$

This technique of comparing very short segments of curves to straight line segments is a powerful and important one in calculus; it is used several times in this lecture.

$$\lim_{\theta \to 0} \frac{1 - \cos\theta}{\theta} = 0$$

While calculating the derivatives of $\cos(x)$ and $\sin(x)$, Professor Jerison said that $\lim_{\theta \to 0} \dfrac{1 - \cos\theta}{\theta} = 0$. This is true, but in order to be certain that our derivative formulas are correct we should understand *why* it's true.

As in the discussion of $\sin(\theta)/\theta$, our explanation involves looking at a diagram of the unit circle and comparing an arc with length θ to a straight line segment. (Remember that θ is measured in radians!) As shown in Figure 1, the vertical distance between the endpoints of the arc is $\cos\theta$, and the horizontal distance between the ends of the arc is $1 - \cos\theta$.

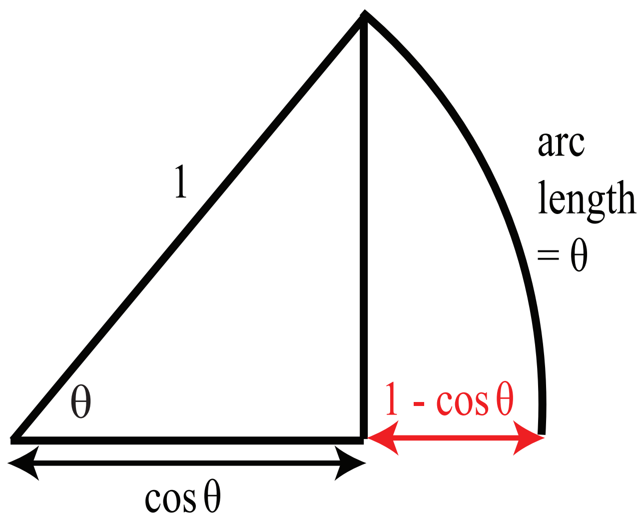

Figure 1: Same figure as for $\frac{\sin x}{x}$ except that the horizontal distance between the edge of the triangle and the perimeter of the circle is marked

From Fig. 2 we can see that as $\theta \to 0$, the horizontal distance $1 - \cos\theta$ between endpoints of the arc (what Professor Jerison calls "the gap") gets much smaller than the length θ of the arc. Hence, $\dfrac{1 - \cos\theta}{\theta} \to 0$.

If you find this hard to believe it may be helpful to use a calculator to verify that if x is small, $1 - \cos x$ is much smaller. You might also study the graph of $y = 1 - \cos x$ near $x = 0$ or use a web application to compare the distance $1 - \cos\theta$ to the arc length θ for very small angles θ.

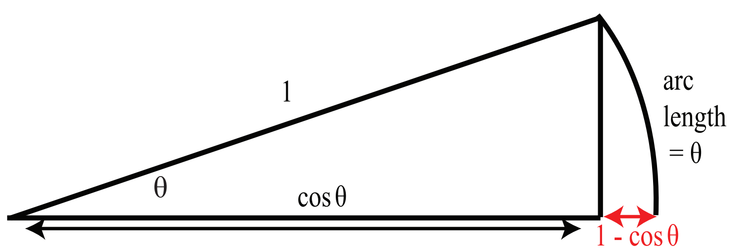

Figure 2: The sector in Fig. 1 as θ becomes very small

Questions and Answers

It's difficult to visualize the relationship between the arc length θ and the segment length $1 - \cos(\theta)$ in the geometric argument that:

$$\lim_{x \to 0} \frac{1 - \cos(\theta)}{\theta} = 0.$$

Professor Jerison spends ten minutes answering student questions and clarifying his argument.

Doesn't θ also tend to 0?

Yes. Whenever we take a derivative we're dividing by a quantity that tends to zero. The quantity $\lim_{x\to 0} \frac{1-\cos(\theta)}{\theta}$ happens to be the derivative of $\cos\theta$ evaluated at zero. It should not surprise us that this ratio tends to zero divided by zero. What we're interested in is that as θ goes to zero the straight line approximating the arc with length θ becomes much longer than the "gap" of length $1 - \cos\theta$.

As the angle θ decreases and we "zoom in", it turns out that the ratio between the size of the "gap" and the length of the "bowstring" gets smaller and smaller – the gap shrinks faster than the bowstring. So, as the measure of angle θ approaches 0, the ratio $\frac{1-\cos(\theta)}{\theta}$ also approaches 0.

In the example of $\dfrac{\sin\theta}{\theta}$ we were comparing two quantities that were about the same size and which approached zero at approximately the same rate, so their ratio approached 1.

Where is the Stata Center?

The Stata Center is a building at MIT that is approximately 100 meters away from the hall in which this lecture is given.

Is $\frac{\cos\theta-1}{\theta}$ the same as $\frac{1-\cos(\theta)}{\theta}$?

$\dfrac{\cos\theta - 1}{\theta}$ is the negative of $\dfrac{1-\cos(\theta)}{\theta}$.

$$-\frac{1-\cos(\theta)}{\theta} = \frac{-1+\cos\theta}{\theta} = \frac{\cos\theta - 1}{\theta}$$

What do you mean by "arc length"?

The arc in question is the portion of the unit circle swept out by the angle θ. Its arc length is its length.

Why is its length θ?

Because we are measuring the angle θ in radians and because the radius of the circle has length 1, the length of the arc swept out by θ will equal the angle θ.

This is especially important in the discussion of $\frac{\sin\theta}{\theta}$. If we were measuring θ in degrees, for example, we would have to compute the arc length before we could compare it to the vertical distance $\sin\theta$. The arc length only equals θ when θ is measured in radians.

The best way to measure θ in this calculation is by length along the unit circle, which is what radians are. Unfortunately, we won't see a proof of this until near the end of the course.

Why is the length of the radius equal to 1?

When working with trig functions like $\cos(x)$ it's extremely useful to think about a circle with radius 1 centered at the point $(0,0)$ because a point (x,y) at angle θ on the circle will have coordinates $(\cos\theta, \sin\theta)$.

The radius shown on the blackboard looks like it's more than one unit long because we "zoomed in" on it as θ became too small to see.

I'm having trouble visualizing how $\frac{1-\cos(\theta)}{\theta} \to 0$.

You are probably not the only one.

First, visualize angle θ getting smaller and smaller.

When θ is very, very small, we can't see what's going on unless we "blow up" or "zoom in on" the picture. (In fact, if we let θ decrease all the way to 0, the triangle collapses and $\frac{1-\cos(\theta)}{\theta} = \frac{0}{0}$. That's why we have to use a limit.)

To estimate the value of $\frac{1-\cos(\theta)}{\theta}$ as θ shrinks, we need to look at geometric interpretations of both the numerator $(1 - \cos\theta)$ and the denominator θ and compare them.

The numerator is the length of the tiny "gap" and the denominator is half the length of the "bow string". From the picture, we see that the gap is much shorter than the bow string when θ is small.

As θ shrinks, we zoom in further, the bow of the circular arc gets closer and closer to the bow string, and so the gap gets smaller and smaller while the bow string appears to stay the same length. So the ratio of the size of the gap to the length of the bow string gets smaller and smaller, approaching zero.

Not only does the size of the gap go to zero, it goes to zero faster than the length of the bow string does.

A geometric proof that the derivative of $\sin x$ is $\cos x$.

At the start of the lecture we saw an algebraic proof that the derivative of $\sin x$ is $\cos x$. While this proof was perfectly valid, it was somewhat abstract – it did not make use of the definition of the sine function.

The proof that $\lim_{\theta\to 0} \frac{\sin\theta}{\theta} = 1$ did use the unit circle definition of the sine of an angle. It also showed that when $x = 0$ the derivative of $\sin x$ is 1:

$$\begin{aligned} \frac{d}{dx}\sin x\Big|_{x=0} &= \lim_{\Delta x\to 0} \frac{\sin(0+\Delta x) - \sin 0}{\Delta x} \\ &= \lim_{\Delta x\to 0} \frac{\sin\Delta x - 0}{\Delta x} \\ &= \lim_{\Delta x\to 0} \frac{\sin\Delta x}{\Delta x} \\ &= 1. \end{aligned}$$

We'll now prove that the derivative of $\sin\theta$ is $\cos\theta$ directly from the definition of the sine function as the ratio $\frac{|\text{opposite}|}{|\text{hypotenuse}|}$ of the side lengths of a right triangle.

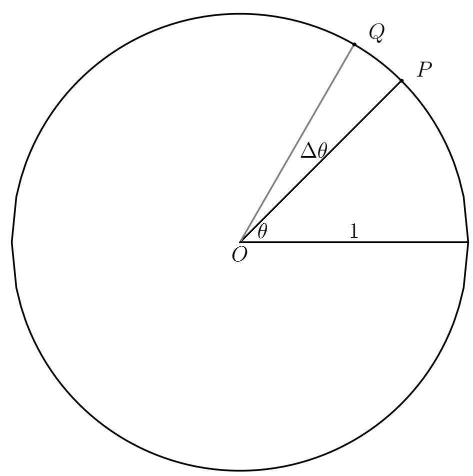

Figure 1: Point P has vertical position $\sin\theta$.

We start with a point P on the unit circle centered at O and the angle θ associated with P. As indicated in Figure 1, $\sin\theta$ is the vertical distance between P and the x-axis. Next, we add a small amount $\Delta\theta$ to angle θ; let Q be the point on the unit circle at angle $\theta + \Delta\theta$. The y-coordinate of Q is $\sin(\theta + \Delta\theta)$. To find the rate of change of $\sin\theta$ with respect to θ we just need to find the rate of change of $y = \sin\theta$.

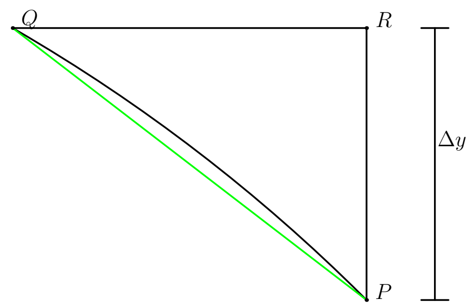

Figure 2: When $\Delta\theta$ is small, $\overset{\frown}{PQ} \approx \overline{PQ}$. Find $\frac{dy}{d\theta}$.

As shown in Figure 2, $\Delta y = |PR|$ and segment PQ is a straight line approximation of the circular arc PQ. If $\Delta\theta$ is small enough, segment PQ and arc PQ are practically the same, so $|PQ| \approx \Delta\theta$.

We're trying to find Δy. Since we know the length of the hypotenuse PQ, all we need is the measure of $\angle QPR$ to solve for $\Delta y = |PR|$.

Since $\Delta\theta$ is small, segment PQ is (nearly) tangent to the circle, and so angle $\angle OPQ$ is (nearly) a right angle. We know that PR is vertical, we know that θ is the angle OP makes with the horizontal, and we can combine these facts to prove that $\angle RPQ$ and θ are (nearly) congruent angles.[1]

The arc length $\Delta\theta$ is approximately equal to the length $|PR|$ of the hypotenuse and angle RPQ is approximately equal to θ. By the definition of the cosine function we get $\cos\theta \approx \frac{|PR|}{\Delta\theta}$. But $|PR|$ is just the vertical distance between Q and P, which is just the difference between $\sin(\theta + \Delta\theta)$ and $\sin\theta$. In other words, when $\Delta\theta$ is very small,

$$\cos\theta \approx \frac{\sin(\theta + \Delta\theta) - \sin\theta}{\Delta\theta}.$$

As $\Delta\theta$ approaches 0, segment QP gets closer and closer to arc QP and angle QPO gets closer and closer to a right angle, so the value of $\frac{(\sin(\theta+\Delta\theta)-\sin\theta)}{\Delta\theta}$ gets closer and closer to $\cos\theta$. We conclude that:

$$\lim_{\Delta\theta\to 0} \frac{\sin(\theta + \Delta\theta) - \sin\theta}{\Delta\theta} = \cos\theta$$

and thus that the derivative of $\sin\theta$ is $\cos\theta$.

[1] Professor Jerison does this by rotating and translating angle θ to coincide with angle RPQ. Another way to see this is to extend segment RP until it intersects the horizontal line through O at point S, then note that $m\angle RPQ + m\angle QPO + m\angle OPS = \pi$ and also $\theta + m\angle PSO + m\angle OPS = \pi$. Since $m\angle QPO \cong m\angle PSO$, we get $m\angle RPQ \cong \theta$. (If $\theta > \pi/2$ a different, but similar, argument applies.)

General Derivative Rules

We've just seen some specific rules for taking the derivatives of the cosine and sine functions. Here are some general rules which we'll discuss in more detail later.

Product Rule

$$(uv)' = u'v + uv'$$

The way that you should remember this is by thinking about changing one function (u or v) at a time. This is a good general procedure when taking derivatives involving multiple functions.

Quotient Rule

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2} \quad (v \neq 0).$$

You'll see proofs of these soon, and you should be able to prove facts like these for your homework and exams.

Introduction to General Rules for Differentiation

We've now seen several specific rules for differentiation; for example, x^n is nx^{n-1}. We've even seen a few examples using this formula. We've also seen some general rules for extending these calculations. For instance, $(cu)' = c \cdot u'$ and $(u + v)' = u' + v'$.

Today we'll learn more general rules; how to differentiate a product of functions, a quotient of functions, and best of all a composition of functions. At the end we'll learn something about higher derivatives.

Product formula (General)

The product rule tells us how to take the derivative of the product of two functions:

$$(uv)' = u'v + uv'$$

This seems odd — that the product of the derivatives is a sum, rather than just a product of derivatives — but in a minute we'll see why this happens.

First, we'll use this rule to take the derivative of the product $x^n \sin x$ – a function we would not be able to differentiate without this rule. Here the first function, u is x^n and the second function v is $\sin x$. According to the specific rule for the derivative of x^n, the derivative u' must be nx^{n-1}. If $v = \sin x$ then $v' = \cos x$. The product rule tells us that $(uv)' = u'v + uv'$, so

$$\frac{d}{dx} x^n \sin x = nx^{n-1} \sin x + x^n \cos x.$$

By applying this rule repeatedly, we can find derivatives of more complicated products:

$$\begin{aligned} (uvw)' &= u'(vw) + u(vw)' \\ &= u'vw + u(v'w + vw') \\ &= u'vw + uv'w + uvw'. \end{aligned}$$

Now let's see why this is true:

$$\begin{aligned} (uv)' &= \lim_{\Delta x \to 0} \frac{(uv)(x + \Delta x) - (uv)(x)}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{u(x + \Delta x)v(x + \Delta x) - u(x)v(x)}{\Delta x} \end{aligned}$$

We want our final formula to appear in terms of u, v, u' and v'. We know that $u' = \lim_{\Delta x \to 0} \frac{u(x + \Delta x) - u(x)}{\Delta x}$, and we see that $u(x + \Delta x)v(x + \Delta x) - u(x)v(x)$ looks a little bit like $(u(x + \Delta x) - u(x))v(x)$. By using a little bit of algebra we can get $(u(x + \Delta x) - u(x))v(x)$ to appear in our formula; this process is described below.

First, notice that:

$$u(x + \Delta x)v(x) - u(x + \Delta x)v(x) = 0.$$

Adding zero to the numerator doesn't change the value of our expression, so:

$$(uv)' = \lim_{\Delta x \to 0} \frac{u(x + \Delta x)v(x) - u(x)v(x) + u(x + \Delta x)v(x + \Delta x) - u(x + \Delta x)v(x)}{\Delta x}.$$

We then re-arrange that expression to get:

$$(uv)' = \lim_{\Delta x \to 0} \left[\left(\frac{u(x + \Delta x) - u(x)}{\Delta x} \right) v(x) + u(x + \Delta x) \left(\frac{v(x + \Delta x) - v(x)}{\Delta x} \right) \right]$$

We proved that if u and v are differentiable they must be continuous, so the limit of the sum is the sum of the limits:

$$\left[\lim_{\Delta x \to 0} \frac{u(x+\Delta x)-u(x)}{\Delta x}\right] v(x) + \lim_{\Delta x \to 0} \left(u(x+\Delta x) \left[\frac{v(x+\Delta x)-v(x)}{\Delta x}\right]\right)$$

or in other words,

$$(uv)' = u'(x)v(x) + u(x)v'(x).$$

Note: we also used the fact that:

$$\lim_{\Delta x \to 0} u(x+\Delta x) = u(x),$$

which is true because u is differentiable and therefore continuous.

Quotient Rule

Now that we know the product rule we can find the derivatives of many more functions than we used to be able to. Our next step toward "differentiating everything" will be to learn a formula for differentiating quotients (fractions). The rule is:

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}$$

Why is this true? The definition of the derivative tells us that:

$$\left(\frac{u}{v}\right)' = \lim_{\Delta x \to 0} \frac{\dfrac{u(x+\Delta x)}{v(x+\Delta x)} - \dfrac{u(x)}{v(x)}}{\Delta x}$$

This is an unwieldy expression. We'll start to make sense of it by simplifying the numerator and by creating two new variables $\Delta u = u(x + \Delta x) - u(x)$ and $\Delta v = v(x + \Delta x) - v(x)$.

$$\begin{aligned}
\frac{u(x+\Delta x)}{v(x+\Delta x)} - \frac{u(x)}{v(x)} &= \frac{u+\Delta u}{v+\Delta v} - \frac{u}{v} \quad \text{(since } u(x+\Delta x) = u(x) + \Delta u\text{)} \\
&= \frac{(u+\Delta u)v - u(v+\Delta v)}{(v+\Delta v)v} \quad \text{(common denominator)} \\
&= \frac{uv + (\Delta u)v - uv + u(\Delta v)}{(v+\Delta v)v} \quad \text{(distribute } u \text{ and } v\text{)} \\
&= \frac{(\Delta u)v - u(\Delta v)}{(v+\Delta v)v} \quad \text{(because } uv - uv = 0\text{)}
\end{aligned}$$

Now that we've simplified the numerator, we can use it to simplify the difference quotient:

$$\begin{aligned}
\frac{\dfrac{u(x+\Delta x)}{v(x+\Delta x)} - \dfrac{u(x)}{v(x)}}{\Delta x} &= \frac{\dfrac{(\Delta u)v - u(\Delta v)}{(v+\Delta v)v}}{\Delta x} \\
&= \frac{1}{\Delta x}\,\frac{(\Delta u)v - u(\Delta v)}{(v+\Delta v)v} \\
&= \frac{\left(\dfrac{\Delta u}{\Delta x}\right)v - u\left(\dfrac{\Delta v}{\Delta x}\right)}{(v+\Delta v)v}
\end{aligned}$$

we're assuming that v is differentiable and therefore continuous, so $\lim\limits_{x \to 0} v(x + \Delta x) = v(x)$. Hence, by the definition of the derivative,

$$\frac{\left(\dfrac{\Delta u}{\Delta x}\right)v - u\left(\dfrac{\Delta v}{\Delta x}\right)}{(v+\Delta v)v} \quad \longrightarrow \quad \frac{v\left(\dfrac{du}{dx}\right) - u\left(\dfrac{dv}{dx}\right)}{v^2} \qquad \text{as } \Delta x \to 0.$$

We conclude that:

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}.$$

Example: Reciprocals

Let's use the quotient rule in a simple example. The quotient rule tells us that:

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{\frac{du}{dx}v - u\frac{dv}{dx}}{v^2}$$

In this example u will be 1, so we'll be finding the derivative of $\frac{1}{v}$, the reciprocal of v.

$$\frac{d}{dx}\left(\frac{1}{v}\right) = ?$$

We're going to use the formula above. We know $u = 1$ and $v = v$, so we still need to find $\frac{du}{dx}$ and $\frac{dv}{dx}$ before we can apply the formula.

The derivative of a constant (like 1) is zero, so $\frac{du}{dx} = 0$. We don't know what v is, so we'll just write $\frac{dv}{dx} = v'$. Plugging all this in to the quotient rule formula we get:

$$\begin{aligned}\frac{d}{dx}\left(\frac{1}{v}\right) &= \frac{0 \cdot v - 1v'}{v^2} \\ &= \frac{-v'}{v^2} \\ &= -v^{-2}v'\end{aligned}$$

Now we have a general formula that lets us differentiate reciprocals! Next, let's use this formula to see what happens when $u = 1$ and $v = x^n$. Here again $\frac{du}{dx} = 0$ and now $v' = \frac{d}{dx}x^n = nx^{n-1}$.

$$\begin{aligned}\frac{d}{dx}\left(\frac{1}{x^n}\right) &= -v^{-2}v' \\ &= -(x^n)^{-2}(nx^{n-1}) \\ &= -x^{-2n}(nx^{n-1}) \\ &= -nx^{-n-1}\end{aligned}$$

But $\frac{1}{x^n} = x^{-n}$, which is x to a power. We have a rule for taking the derivative of x to a positive power; how does that compare to our new rule for the derivative of x to a negative power?

$$\frac{d}{dx}x^{-n} = -nx^{-n-1}$$

This agrees with the formula $\frac{d}{dx}x^n = nx^{n-1}$, so the quotient rule confirms that our rule for taking the derivative of x^n works even when n is negative.

Chain Rule

The product rule tells us how to find the derivative of a product of functions like $f(x) \cdot g(x)$. The composition or "chain" rule tells us how to find the derivative of a composition of functions like $f(g(x))$. Composition of functions is about substitution – you substitute a value for x into the formula for g, then you substitute the result into the formula for f. An example of a composition of two functions is $y = (\sin t)^{10}$ (which is usually written as $y = \sin^{10} t$).

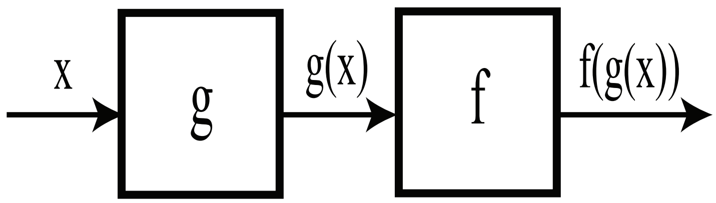

Figure 1: Composition of functions: $(f \circ g)(x) = f(g(x))$

One way to think about composition of functions is to use new variable names. For example, for the function $y = \sin^{10} t$ we can say $x = \sin t$ and then $y = x^{10}$. Notice that if you plug $x = \sin t$ in to the formula for y you get back to $y = \sin^{10} t$. It's good practice to introduce new variables when they're convenient, and this is one place where it's very convenient.

So, how do we find the derivative of a composition of functions? We're trying to find the slope of a tangent line; to do this we take a limit of slopes $\frac{\Delta y}{\Delta t}$ of secant lines. Here y is a function of x, x is a function of t, and we want to know how y changes with respect to the original variable t. Here again using that intermediate variable x is a big help:

$$\frac{\Delta y}{\Delta t} = \frac{\Delta y}{\Delta x} \cdot \frac{\Delta x}{\Delta t}$$

because when we perform the multiplication, the small change Δx cancels.

The derivative of y with respect to t is $\lim\limits_{\Delta t \to 0} \dfrac{\Delta y}{\Delta t}$; what happens when Δt gets small? Because $x = \sin t$ is a continuous function, as Δt approaches 0, Δx also approaches zero. It turns out that:

$$\lim_{\Delta t \to 0} \frac{\Delta y}{\Delta t} = \frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt} \qquad \longleftarrow \textbf{The Chain Rule!}$$

The derivative of a composition of functions is a product. In the example $y = (\sin t)^{10}$, we have the "inside function" $x = \sin t$ and the "outside function" $y = x^{10}$. The chain rule tells us to take the derivative of y with respect to x and multiply it by the derivative of x with respect to t.

The derivative of $y = x^{10}$ is $\dfrac{dy}{dx} = 10x^9$. The derivative of $x = \sin t$ is $\dfrac{dx}{dt} = \cos t$. The chain rule tells us that $\frac{d}{dt} \sin^{10} t = 10x^9 \cdot \cos t$. This is correct, but if a friend asked you for the derivative of $\sin^{10} t$ and you answered $10x^9 \cdot \cos t$ your friend wouldn't know what x stood for. The last step in this process is to rewrite x in terms of t:

$$\frac{d}{dt} \sin^{10} t = 10(\sin t)^9 \cdot \cos t = 10 \sin^9 t \cdot \cos t.$$

Here is another way of writing the chain rule:

$$\frac{d}{dx}(f \circ g)(x) = \frac{d}{dx} f(g(x)) = f'(g(x))g'(x).$$

Example: $\sin(10t)$

The chain rule (composition rule) says that $\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt}$. In other words, the derivative of the composition of functions $f(g(t))$ is the derivative of the outside function $f(x)$ times the derivative of the inside function $g(t)$.

For the example $\frac{d}{dt}\sin(10t)$, the inside function is $x = 10t$ and the outside function is $y = \sin x$. Using the rules we know, we can compute that $\frac{dy}{dx} = \cos x$ and $\frac{dx}{dt} = 10$, so:

$$\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt} = \cos x \cdot 10.$$

Since we're the only ones who know the value of x in this formula, we replace x by $10t$ to get:

$$\frac{dy}{dt} = \cos(10t) \cdot 10 = 10\cos(10t).$$

Once you've had more practice using the chain rule, you won't always need the variable x that represents the inside function. When you look at $\frac{d}{dt}\sin(10t)$ you might say to yourself: "The derivative of the outside function, sine, is cosine. I'm plugging $10t$ into it. And the derivative of $10t$ is just 10. So $\frac{d}{dt}\sin(10t) = \cos(10t) \cdot 10$.

Higher Derivatives

Higher derivatives are derivatives of derivatives. Given a differentiable function $u = u(x)$ its derivative u' is a new function, which we may be able to differentiate again to get $(u')' = u''$.

For example, if $u(x) = \sin x$ then $u' = \cos x$ and $u'' = -\sin x$. We can go on: $(u'')' = u''' = -\cos x$ ($u''' = u^{(3)}$ is called the third derivative of u and u'' is the second derivative) and $u'''' = u^{(4)} = \sin x$. The function $\sin x$ is a special example – we won't usually "come back to" the function we started with.

Since there's more than one way to write derivatives, there's more than one notation for higher derivatives.

Notations

$f'(x)$	Df	$\frac{df}{dx}$	$\frac{d}{dx}f$
$f''(x)$	D^2f	$\frac{d^2f}{dx^2}$	$\left(\frac{d}{dx}\right)^2 f$
$f'''(x)$	D^3f	$\frac{d^3f}{dx^3}$	$\left(\frac{d}{dx}\right)^3 f$
$f^{(n)}(x)$	D^nf	$\frac{d^nf}{dx^n}$	$\left(\frac{d}{dx}\right)^{\text{n}} f$

The symbols D and $\frac{d}{dx}$ represent "operators" which can be applied to a function. When you apply one of these operators to a function you get the derivative of that function.

Example: $D^n x^n$

Let's calculate the n^{th} derivative of x^n

$$D^n x^n = ? \quad (n = 1, 2, 3, ...)$$

Let's start small and look for a pattern:

$$\begin{aligned} Dx^n &= nx^{n-1} \\ D^2x^n &= n(n-1)x^{n-2} \\ D^3x^n &= n(n-1)(n-2)x^{n-3} \\ &\vdots \\ D^{n-1}x^n &= (n(n-1)(n-2)\cdots 2)x^1 \end{aligned}$$

We can guess this $(n-1)^{st}$ derivative from the pattern established by the first three derivatives. The power of x decreases by 1 at every step, so the power of x on the $(n-1)^{st}$ step will be 1. At each step we multiply the derivative by the power of x from the previous step, so at the $(n-1)^{st}$ step we'll be multiplying by the previous power 2 of x.

Differentiating one more time we get:

$$D^n x^n = (n(n-1)(n-2)\cdots 2 \cdot 1)1$$

The number $(n(n-1)(n-2)\cdots 2\cdot 1)$ is written $n!$ and is called "n factorial". What we've just seen forms the basis of a proof by mathematical induction that $D^n x^n = n!$. So $D^n x^n$ is a constant!

The final question for the lecture is: what is $D^{n+1}x^n$?

Answer: It's the derivative of a constant, so it's 0.