Lecture 3

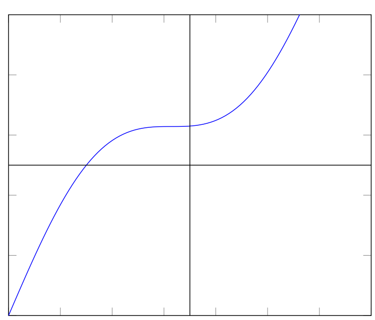
Derivatives, and derivatives of polynomials

We'll start with some review to get ourselves situated. The new material for today starts in Section 3.2.

3.1 Review

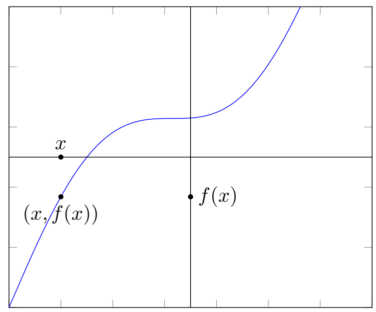
We've been using some notation like $f(x)$, and talking about graphs. Let's make some of these ideas explicit. This also serves as a review of what you've learned in some past classes.

Below is the graph of a function f.



As you may have learned in a previous class, a choice of a number, x, determines a point on the graph! The point is determined by moving x units along the x-axis, and then moving $f(x)$ units vertically. The symbol we use for this point is:

$$(x, f(x)).$$



In the above picture, x is some negative number, and $f(x)$ happens also to be some negative number. Moreover, because f is a function (so its graph passes the vertical line test) if P is a point on the graph of f, we may write the coordinates of P as

$$P = (x, f(x))$$

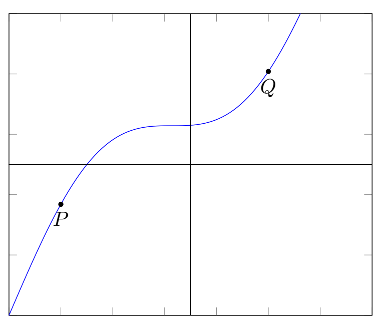
for some number x.

Because a number x determines a point P on the graph of f, and because the point P determines x, we will often talk about **how a function behaves "at x"** instead of how the function behaves at P.

3.1.1 Slopes of secant lines

We talked last time about secant lines, too. So let's review the notation.

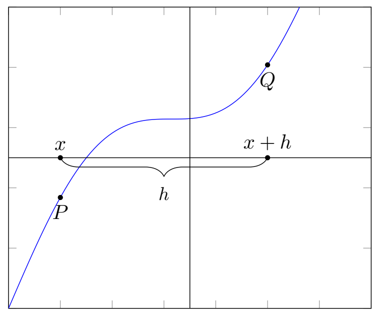
Let's say somebody chooses another point Q on the graph of f.



We may ask about the *horizontal* difference between P and Q—that is, what is the difference between the x-coordinate of P, and the x-coordinate of Q? Whatever it is, let us call it h:

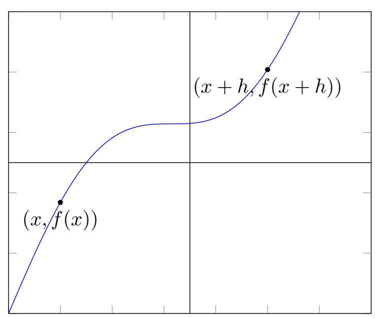
$$h = \text{The difference between the } x\text{-coordinates of } P \text{ and } Q.$$

So that the x-coordinate of Q is given by $x + h$.



(Note that h could be a positive or a negative number. In our pictures, h happens to be positive.) Then the coordinates of Q are given by:

$$Q = (x + h, f(x + h)).$$



As we saw last time:

Proposition 3.1.1. the slope of the secant line through P and Q is given by the formula

$$\frac{f(x + h) - f(x)}{h}.$$

Remark 3.1.2. In these notes, you will see me label sections by "Remarks" or "Propositions." (There will be other labels you'll see as the class goes on.)

A "Remark" is just a comment I would like to make.

A "Proposition" is a term used throughout mathematics. A Proposition is a statement that is true, and useful for getting an idea for what's going on, and also not too difficult to convince somebody of. For many students, you won't lose too much sleep if you replace the word "Proposition" by the word "Fact," but it is also very healthy to wonder *why* certain facts are true. If the fact is a Proposition, I promise you will be able to see why the fact is true if you spend a reasonable amount of time thinking it through.

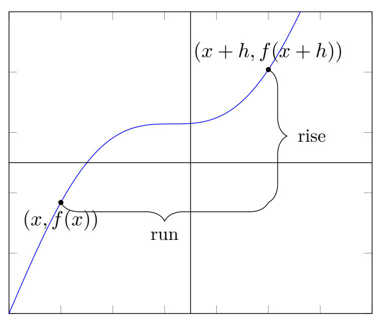
A "Proof" is a series of sentences of equations to convince you that the Proposition is true.

Proof. The slope of the secant line is given by "rise over run." The rise is the difference in the y-coordinates of P and Q, so

$$\text{rise} = f(x+h) - f(x).$$

The run is the difference in the x-coordinates of P and Q, so

$$\text{run} = (x+h) - x = h.$$



Then

$$\frac{\text{rise}}{\text{run}} = \frac{f(x+h) - f(x)}{h}.$$

□

Remark 3.1.3. The above proposition is useful because it takes a *geometric* idea (like drawing lines through points of graphs) and converts it into *algebra* (e.g., a formula).

For example, even if you had no idea what the graph of f looked like, if somebody gives you a formula for f, you can now compute the slope of a secant line. This is very powerful. It takes a lot more time to try to visually draw and measure something, than to just compute it using algebra.

Because we'll see this fraction a lot, we gave it a name: $\frac{f(x+h)-f(x)}{h}$ is called a *difference quotient*. You became familiar with this expression for today's lecture.

3.1.2 The slopes of tangent lines

So let's choose a function f and a point P on the graph of f.

Remember that the tangent line to f at P is the line that the secant lines through P and Q approach as Q gets closer and closer to P. Well, if a bunch of lines approach a single line, then it stands to reason that the slopes of the those lines approach the slope of the single line.

Well, how can we talk about what it means for "Q to approach P"? Remember that, earlier, we chose h to be the horizontal difference between P and Q. So if the point Q is approaching P, then surely h is shrinking to zero!

On the other hand, we saw that the fraction $\frac{f(x+h)-f(x)}{h}$ is the slope of the secant line between P and Q.

So what we need to understand is the following:

Question 3.1.4. How does the number $\frac{f(x+h)-f(x)}{h}$ behave as h becomes closer and closer to zero?

Warning 3.1.5. Note that the fraction above has h in the denominator. In other words, *you cannot just plug in* $h = 0$ into the fraction. (One of the golden rules of mathematics is: You cannot divide by zero.)

3.1.3 The derivative (using words)

This question above (Question 3.1.4) is at the heart of calculus.[1]

Definition 3.1.6. Let f be a function, and x a number. Then *the derivative of f at x* is the number that $\frac{f(x+h)-f(x)}{h}$ approaches as h gets closer and closer to zero (**if such a number exists**).

Remark 3.1.7 (What is a definition?)**.** A *definition* is, as you know, something you usually find in a dictionary. In a set of math notes, or in a math textbook, a definition is like a shortcut.

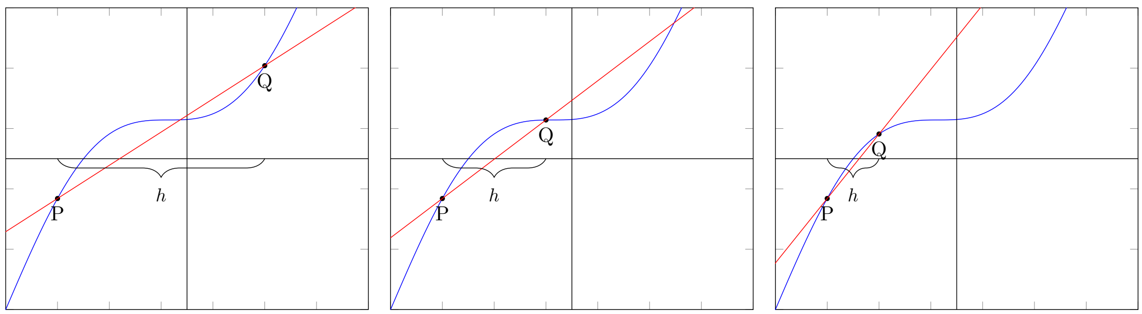
For example, the above definition of a derivative is very wordy! Too many words. So instead of saying all that ("the number that this fraction approaches as h goes to zero," if it exists) we get to just say "the derivative." Isn't that a lot shorter?

For example, I could say "Erica is my friend," rather than saying "Erica is someone with whom I have a good relationship and with whom I sometimes hang out." If a mathematician were to introduce the word *friend* in a textbook, they might write:

[1]Really, the heart of half of calculus. The other half of calculus is devoted to integrals, which we'll see later on this semester.

Definition. Let P be a person. Then P is called a *friend* if P is a person with whom you have a good relationship and with whom you sometimes hang out.

Just as a visual reminder, here is what's geometrically happening as "h approaches zero":



Because h is the difference in the x-coordinates of P and Q, as h approaches zero, P and Q are getting closer and closer. And remember, we are not moving P around; we're just letting Q move. So "as h approaches zero," is a way to say "as Q moves closer to P."

3.1.4 The derivative (using limit notation)

Our definition of limit is a mouthful. That's why we also invent notation to make things easier. For example, $f(x) = 5x^3 + 3x + 1$ if far easier to see than "let f be a function that takes a number, cubes it, multiplies the result by 5, then takes the original number, multiplies it by 3, then adds that, then finally adds 1 to all that previous stuff." So here is the notation we will be using in this class:

Definition 3.1.8 (The derivative, using limit notation)**.** Let f be a function and x a number. Then the *derivative of f at x* is the number

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h}$$

if it exists.

This notation is read "the limit as h approaches 0 of $\frac{f(x+h)-f(x)}{h}$." This is just notation for now, and you're not expected to have a deep understanding of this notation yet. We'll get more intimate with this notation later in the semester.

3.1.5 Example

Let's choose $f(x) = x^2 + 10$, and $x = 3$. Then

$$\frac{f(3+h) - f(3)}{h} = \frac{((3+h)^2 + 10) - (3^2 + 10)}{h} = \frac{3^2 + 2 \cdot 3h + h^2 + 10 - 3^2 - 10}{h}$$
$$= \frac{6h + h^2}{h}.$$

Now, the question is, how does this fraction behave as h approaches zero? Here's the thing: Whenever h does not equal zero, the above fraction can be simplified:

$$\frac{6h + h^2}{h} = \frac{6h}{h} + \frac{h^2}{h} = 6 + h.$$

So, now we ask:

Question. What does $6 + h$ become as h approaches zero?

Well, as h gets smaller and smaller, $6 + h$ becomes a number closer and closer to 6. So the answer to the question is: $6 + h$ becomes 6 as h approaches zero.

This is our first derivative that we've ever computed! The derivative of $f(x) = x^2 + 10$ at $x = 3$ is given by 6.

3.1.6 Summary

Whenever you've solved a problem, it's good to look back on what you did.

1. We first wrote out the difference quotient $\frac{f(x+h)-f(x)}{h}$ and plugged in $x = 3$.
2. We simplified it as far we could, keeping in mind that we *should not* divide by h. We ended up with $\frac{6h+h^2}{h}$.
3. We then tried to understand the behavior of the fraction when h does not equal zero. This is a great thing to do, because when h does not equal zero, we *can* simplify the fraction when the numerator has h in every term. We ended up with having to understand how $6 + h$ behaves as h equals zero.
4. Well, we can reason out that as h approaches zero, $6 + h$ approaches 6. And that's our answer.

The notation we can use for this is:

$$\lim_{h \to 0} \frac{f(3+h) - f(3)}{h} = 6$$

.

3.2 Notations and language for the derivative

Because the derivative is something we'll be using so much, we're going to want some briefer notation. Here are the most common notations used for derivatives:

Notation 3.2.1. Let f be a function and x a number. Then the derivative at x is denoted by any of the following:

- $f'(x)$
- $\frac{df}{dx}(x)$
- $\left(\frac{d}{dx}f\right)(x)$.

The above notations all mean the same thing.

The notation we'll use most often is $f'(x)$. This is read "f **prime of** x."

The notation $\frac{df}{dx}$ is read "**dee eff dee ex.**"

The notation $\frac{d}{dx}f$ is read "**dee dee ex of f.**"

Sometimes, we will give functions different names – instead of f, we may use a letter like g, or h, or A. And, we may also use a different letter for the input variable – instead of x, we may use t (especially if the function takes time as an input), or s.

The notation f' doesn't tell you what the name of the input variable is. But the notation $\frac{df}{dt}$ makes it clear that the input variable is t. So does the notatoin $\frac{d}{dt}f$. When we want to be clear about the name of the input variable, we say that $\frac{df}{dt}$ is the derivative of f **with respect to** t.

For example, we say that $\frac{df}{dx}$ is the derivative of f with respect to x.

Example 3.2.2. In lab you saw that if $f(x) = 3x + 2$, then the derivative of f at 2 is given by 3. In other words,

$$f'(2) = 3.$$

You also saw in lab that if $g(x) = 5x^2$, then the derivative of g at 2 is given by 20. So

$$g'(2) = 20.$$

(That is, "g prime of 2 is 20.") You could also have written

$$\frac{dg}{dx}(2) = 20.$$

3.3 (Intuition) The derivative as a function

In lab, you probably plugged in a value for x, and then evaluated the difference quotient. For example, for $g(x) = 5x^2$, to evaluate the derivative at $x = 2$, you probably began by writing out the fraction

$$\frac{g(2+h) - g(2)}{h} = \frac{5(2+h)^2 - 5(2)^2}{h}.$$

But one think you could also have done, is to *not* plug in $x = 2$, and just leave it as x. It may feel a little strange to have two variables (both x and h) floating around, but we can still treat them as numbers and perform our calculations:

$$\begin{aligned}\frac{g(x+h) - g(x)}{h} &= \frac{5(x+h)^2 - 5x^2}{h} \\ &= \frac{5x^2 + 10xh + 5h^2 - 5x^2}{h} \\ &= \frac{10xh + 5h^2}{h}.\end{aligned}$$

This final expression, as usual, simplifies when $h \neq 0$ (i.e., when h does not equal zero):

$$10x + 5h.$$

Then, as h approaches zero, the expression $10x + 5h$ becomes $10x$. So we conclude:

The derivative of g at x is given by $10x$.

Or, using the "prime" notation we just learned, we can write:

$$g'(x) = 10x.$$

But wait! This g' is a new function!

What do I mean? Remember that a function is something that eats a number, and spits out another number. For example, when we write $f(x) = 3x$, we mean that f is a function that takes in a number, then spits out three times that number. Likewise, $g(x) = 5x^2$ is a function that has you input a number (called x) and that squares it, then multiplies it by five ($5x^2$).

Likewise, we see that g' is something that will take in a number (called x) and output a new number (called $10x$).

Upshot. If f is a function, then f' is a new function.

Terminology. If f is a function and x is a number, then the number $f'(x)$ is called the derivative of f *at x*. On the other hand, the function f' is called *the derivative of f*.

Remark 3.3.1 (Intuition)**.** So what is up with this "new function" called the derivative of f? Graphs help us think intuitively, especially if we like visuals.

A function f can be represented by a graph, as you know. The new function f' tells us – for every number x – what the *slope of the tangent line* is at the point x (of the original function f).

3.4 Interpretation of derivatives

We have seen that the slope of a tangent line should be interpreted as an instantaneous rate of change. So f' is a function which – at a point x, tells you the rate at which f is changing at x. And $f'(x)$ is that rate.

Example 3.4.1. Suppose f is a function which describes the position of a train at time t – specifically, for every time t, the number $f(t)$ is how far along the train is along a track. Let's assume t is measured in minutes and $f(t)$ is in units of feet.

Then $f'(t)$ naturally has units of feet per minute (remember that $f'(t)$ represents the slope of the tangent line at t, and slopes have units!). Indeed, $f'(t)$ is a function that tells us the speed of the train at time t, in feet per minute.

Example 3.4.2. Suppose $h(x)$ is the total cost (in dollars) of producing x many Tickle Me Elmo dolls (in "dolls"). Though we of course only produce 1 doll, 2 dolls, 3 dolls, and not – say – 4.32 dolls in real life, it is very common to model h using a function that knows how to take in numbers like 4.32.

Then $h'(x)$ approximates the increase in cost associated to producing one more doll when you are already producing x many dolls.

This is called the *marginal cost* in economics and business – and it is rather subtle. Keep in mind that it is often cheaper to make one more doll if you are already going to produce 10,000 dolls, than to produce one more doll if you are only going to produce 10 dolls.

The unit of $h'(x)$ is "dollars per doll."

3.5 Basic derivative laws (computing our first derivatives)

Our ultimate goal in life (just kidding—our goal in this class) is to take a function f, and to be able to understand the new function f'. That is, to understand the derivative of f.

Functions f can be complicated. So it'll be useful to understand how to compute derivatives in the simplest of situations. Indeed, much of the computational[2] content of a calculus class is centered around being given a function f, and being able to compute its derivative.

Here, I'm going to tell you some basic computational tools we can use to compute derivatives. In many textbooks, these tools are often called "laws." I don't like this word, because it makes these tools feel handed down to you by some higher authority (like a textbook or a professor). They are not. They are fundamental truths of nature that you could have discovered yourself.

Here are some of the most basic laws about how to write f' given f:

3.5.1 Derivatives of powers

In lab, we have already computed the derivative of a few functions. Let me summarize some of what we have already computed, and add on some more examples of derivatives, to the following table:

Table 3.1: Derivatives of some functions – for example, the above table tells us that the derivative of x^4 is the function $4x^3$.

$f(x)$	$f'(x)$
x^2	$2x$
x^3	$3x^2$
x^4	$4x^3$
x^5	$5x^4$

I want to emphasize these are not just answers that a professor is spitting out at you. These are answers *you* could compute yourself. To see more details on how, see Section 3.8

You probably see a pattern – try taking a moment to guess what the derivative of x^6 is. How about x^{13}?

Proposition 3.5.1 (Power law). If n is any whole number, the derivative of x^n is the function

$$nx^{n-1}.$$

[2]Remember that in this class, I emphasize three ways to think about every idea – intuition, computation, and interpretation.

In other words, if you have a function called "raise x to the nth power," its derivative is computed by a function which raises x to the $(n-1)$st power – i.e., it raises x to one less power – and also scales the function by n.

Example 3.5.2. If $f(x) = x^{13}$, then $f'(x) = 13x^{12}$.

Example 3.5.3. This example is a really great review of exponents. Remember that *anything raised to the 0th power equals 1.* For example, $5^0 = 1$. And $x^0 = 1$. And $(3x^2 + 9x + \pi)^0 = 1$.

Also remember that *anything raised to the 0th power is itself.* For example, $5^1 = 5$, and $x^1 = x$, and $(3x^2 + 9x + \pi)^1 = 3x^2 + 9x + \pi$.

The power law works when $n = 1$: If $f(x) = x^1 = x$, then $f'(x) = 1x^0 = 1 \cdot 1 = 1$. But we'll see later on why you shouldn't need the power law to compute the derivative of x. There are many, many ways to see that the derivative of x is 1.

Example 3.5.4 ($n = 1$)**.** If $f(x) = x$, then $f'(x) = 1$.

Example 3.5.5 ($n = 2$)**.** If $f(x) = x^2$, then $f'(x) = 2x$.

Example 3.5.6 ($n = 3$)**.** If $f(x) = x^3$, then $f'(x) = 3x^2$.

Example 3.5.7 ($n = 4$)**.** If $f(x) = x^4$, then $f'(x) = 4x^3$.

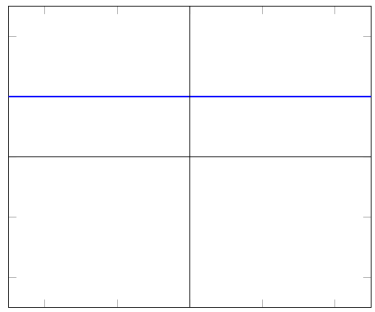
Example 3.5.8 ($n = 5$)**.** If $f(x) = x^5$, then $f'(x) = 5x^4$.

This is a pattern that you'll have to get used to. You will see it *all the time* in this class. Make sure you know the power law!

Remark 3.5.9. I have never found a good physical or visual argument for why the power law is true. The only reason it's true—that I know of—is algebra. (You'll see this in Section 3.8.) This kind of makes sense. Functions of the form x^5 are defined very algebraically. They're not defined using geometric operations like "find the size of an angle," so Mother Nature doesn't give us many immediate ways to think about functions like x^5 geometrically.

3.5.2 Constant functions have zero derivative

Suppose $f(x) = 3$. This is called a "constant function," because no matter what the input number is, f constantly outputs 3. The graph of a constant function is a horizontal line.



(Above, in blue, is an image showing the graph of $f(x) = 3$.) Of course, this line itself has slope zero. In fact, any secant line to f is just the graph of f itself! Accordingly, the tangent line to f is f itself. So f' is another constant function, with value zero.

Law. *If f is a constant function, then $f' = 0$.* That is, regardless of x, $f'(x) = 0$.

Example 3.5.10. If f is the constant function given by $f(x) = \pi$, then f' is the function given by $f'(x) = 0$.

Remark 3.5.11. We can think about this physically, too. If $f(t)$ represents the position of someone at time t, then f being constant means that person isn't moving. So the speed of that person is always zero—that is, the derivative of f is zero.

(Remember that we saw that slopes of lines had to do with speed, when looking at a position-versus-time graph for motion with constant speed.)

3.5.3 Scale a function, scale the derivative

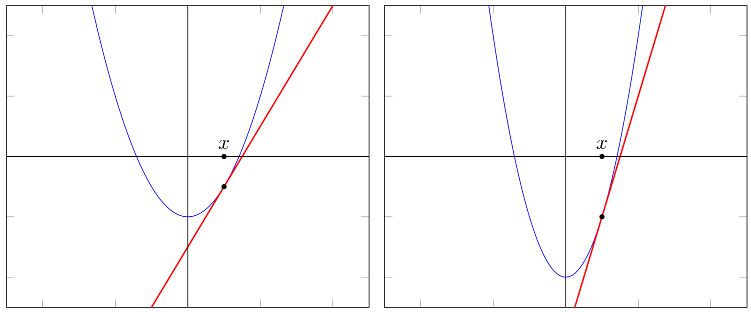
Suppose that f is some function. How is the derivative of f related to the derivative of $5f$? For example, how is the derivative of x^3 related to the derivative of $5x^3$?

Law. $(af)' = a(f')$ (for any number a).

Let's parse this law. It means that if I take a function, multiply it by a number a, and *then* take the derivative, I'll get the same answer as first taking the derivative, then multiplying by a.

Example 3.5.12. Suppose $f(x) = 5x^2$. We saw earlier that f' is the function given by $f'(x) = 10x$. If we let $g(x) = 10x^2 = 2f(x)$, then $g' = 2f'$, so $g'(x) = 2f'(x) = 20x$.

Sometimes, multiplying by a number a is called "scaling by a." So this law says that if you scale a function, you scale its derivative.



Above, on the left is a picture of the graph of a function $f(x) = x^2 - 2$, together with a tangent line at $x = 1$, in red. On the right is the picture of a graph of a function $2f$, that is, $2x^2 - 4$, together with a tangent line at $x = 1$ (the same x as for the lefthand picture). Though it's not obvious from the pictures, the slope of the tangent line on the right is *twice* the slope of the tangent line on the left. In other words, the slope of the tangent line is scaled by the same factor by which the function was scaled.

Remark 3.5.13. We can again think about this physically. Suppose $f(t)$ is a function that again tells you your position of at time t. What would it mean to scale f by a number a? Well, it means that your position is a times as far at any given time. (For example, if $a = 3$, then your position would be triple what it would have been with the unscaled, original function.) What the scaling law says is that if you're always ending up three times as far, you're always moving three times as fast.

Remark 3.5.14. Here is another way to think about the scaling law – visually. Take an image of a graph, and stretch it vertically by a factor of a (as though it were made of rubber). (If a is less than 1, this would involve shrinking, not stretching, the graph vertically.) This will result in stretching whatever tangent line you are interested, also by a factor of a – thus, whatever rise over run you compute will have a "rise" increased by a factor of a, and an unchanged "run." So the slope is multiplied by a factor of a.

3.5.4 Add functions, add derivatives

Law. $(f + g)' = f' + g'$.

That is, if you add two functions, and then take their derivative, you'll get the same answer as taking the derivative of each function, and then adding them.

Remark 3.5.15. This is a natural-looking law, but it's one of the harder ones to justify using pure geometry. It's easier to justify if you think physically.

Suppose that $f(t)$ represents the position of a train at time t.[3] This could be measured, for example, by someone outside the train, observing the train.

Suppose further that there is a cheetah inside the train, and $g(t)$ represents the position of the cheetah "relative to the train" at time t. This is measured, for example, by somebody *inside* the train, observing the cheetah. [4]

The the function $f+g$, which at time t outputs $f(t)+g(t)$, represents the position of the cheetah.[5] This could be measured, for example, by somebody outside the train, with x-ray vision, observing the cheetah.

So how are the speed of the train, the speed of the cheetah relative to the train, and the actual speed of the cheetah, all related?

Well, the actual speed of the cheetah would be the *sum* of the speed of the train with how fast the cheetah is moving relative to the train![6]

Remember that the slope of a line has to do with speed. We saw this for position-versus-time graphs of objects moving with constant speed. And if our position-versus-time graph is a curve, then we can interpret the *derivative* (which is the slope of the tangent line) as the speed that a (highly accurate) speedometer would show at the time.

So this physical "thought experiment" displays one physical argument why the derivative of a sum of functions should be the sum of their derivatives.

3.6 Derivatives of polynomials

In groupwork and in lab, you will practice taking derivatives of polynomials (without using difference quotients anymore).

What is a polynomial? Here are examples:

[3] For example, it could represent how far the center of the train has moved along the track from a departure station.

[4] To make things even more concrete, $g(t)$ could represent how far away the cheetah is from the center of the train, at time t. In any case, if the cheetah is just seated in its passenger seat, $g(t)$ would be constant.

[5] For example, how far the cheetah is from the departure station.

[6] For example, if the train is moving 60 miles per hour to somebody watching the train from outside, but if the cheetah is just sitting still inside the train, then the outside observer with x-ray vision would perceive the cheetah to be moving at 60 miles per hour. On the other hand, if an inside-the-train observer saw the cheetah running 60 miles per hour in the direction opposite the train's front, then to an outside observer with x-ray vision, the cheetah would look stationary, as though it were running on a treadmill.

- 1
- 3
- πx
- $-3x$
- $2x$
- $x+2$
- $9x^2 + \frac{1}{2}x$
- $\pi x^2 + 2x + 9$
- $3x^3 + \sin(1)x^2 + \cos(\pi)x + \pi^3$.

The key thing to note about all the above expressions is that they all look like

$$ax^3 + bx^2 + cx + d$$

where a, b, c, d are arbitrary numbers. (They could be zero, they could be π^3, whatever!) In fact, there's no limit to the power of x in a polynomial—another example is something like

$$x^{10{,}003} + x^2 + 1.$$

The important thing is that each time the symbol x shows up, the exponent of x is some integer that's not negative, and that x isn't inside some other function like sin, or e^x, et cetera.

Remark 3.6.1. In principle, polynomials are supposed to be the "simplest" kinds of functions. For example, if $f(x) = 3x^3 + 2x^2 + x - 9$, you would be able to tell me things like $f(3)$ and $f(-2)$ by hand. You just need to multiply and add a lot.[7]

They key thing to note is that a polynomial function is always a (i) sum of (ii) scaled versions of (iii) x^n for some n.

And the derivative rules from today show us how to deal with sums, with scaling, and with "power" functions like x^n. So we actually now know how to take derivatives of polynomial functions!

[7] On the other hand, if $f(x) = \sin(x)$, you probably would not be able to tell me, by hand, what $\sin(3)$ is. You would need a calculator! We'll see later in this class, if we're lucky, how you might able to calculate $\sin(3)$ to many decimal places without a calculator.

Example 3.6.2. Find the derivative of $f(x) = 3x^2 + x - 9$.

Answer. f is a sum of three terms: $3x^2$, x, and -9. So let's try to take the derivative of each.

The derivative of -9 is easy. This is a constant function, so its derivative is zero.

The derivative of x is also easy. For example, the graph of $g(x) = x$ is just a line with slope 1, so its derivative is always 1. You could also have used the power law if you wanted, because $x = x^1$.

Finally, $3x^2$ is 3 times x^2. We know that the derivative of x^2 is given by $2x$ by the power law. So, scaling this by 3, we conclude that the derivative of $3x^2$ is given by $6x$.

Putting this all together:

$$\begin{aligned}(3x^2 + x - 9)' &= (3x^2)' + (x)' + (-9)' \\ &= 3 \cdot (x^2)' + 1 + 0 \\ &= 3 \cdot 2x + 1 + 0 \\ &= 6x + 1 + 0 \\ &= 6x + 1. \end{aligned} \tag{3.6.1}$$

And that's it!

Example 3.6.3. Find the derivative of $f(x) = x^5 - 7x^4 + 2x + 13$.

I'll tell you that the answer is $5x^4 - 28x^3 + 2$. Can you figure out why?

Example 3.6.4. Find the derivative of $7x^3 - \pi x - 10$ at 8.

In a problem like this, you need to not only compute the derivative (as a function – so involving x) you must also compute its value (at $x = 8$).

We see that the derivative is

$$(7x^3 - \pi x - 10)' = (7x^3)' - (\pi x)' - (10)' = 21x^2 - \pi.$$

Plugging in $x = 8$, we find that the derivative at 8 is

$$21(8)^2 - \pi = 21 \times 64 - \pi = 192 - \pi.$$

So the answer is $192 - \pi$.

Exercise 3.6.5. For each of the following functions, compute the derivative at $x = 3$.

(a) x^3

(b) x^7

(c) 1

(d) x

(e) $5x^2$

(f) $7x$

(g) $8x^3$

(h) $5x + 3$

(i) $7x^{10} - 9x^2 + 1$

(j) $10x^{99} - \cos(9)x^2 + \pi$

(k) $(x - 3)(x + 10)$

(l) $(x - 3)(x + 10)(x - 1)$

(m) $(x - 3)(x + 3)$

3.6.1 What was the difference quotient all for?

You may have learned today that to take the derivative of x^8, you can just write $8x^7$. This is the power law. And that's fantastic.

So you may question: Why did Hiro bother doing this "the complicated way" by having you practice manipulating expressions like $f(x + h) - f(x)$ and dividing by h?

The answer: Because "this complicated way" is what the derivative actually is.

Here's a good analogy. You all know that $8 \times 7 = 56$ – not because you quickly drew out a large rectangular array of dots and counted the fifty-six dots. It's because you memorized, and became familiar with, the multiplication table.

But if you were to explain to a child what multiplication actually *is*, it would be criminal to say "8 times 7 is just 56 and that's the end of the story." It's not the end of the story, nor the beginning. 8 times 7 represents the outcome of a real process – either adding 8 to itself 7 times, or drawing an array of dots and counting the dots, or computing the area of a rectangle with edge lengths 8 and 7. Multiplication actually represents something.

Likewise, I want you to know what the derivative actually represents. In future classes, a professor may only care that you know that the derivative of x^8 is $8x^7$, and you may only care about what the professor will grade you on. But in real life, if you ever use a derivative, you need to know what it actually means. The derviative represents what happens to slopes of secant lines through P and Q as you move Q closer and closer to P. And the difference quotient – along with the process of making h shrink – is exactly the algebra that represents this geometric intuition.

3.7 For next time

For next time, you should be able to take derivatives of polynomials. You'll get plenty of practice in lab.

3.8 Bonus: Computing derivatives of powers

This is not a "required reading" section. It's just to satisfy your itch – how do we know that Table 3.1 is correct?

Let's first do this in a concrete example – we'll compute the derivative of $f(x) = x^n$ when $n = 4$.

Remember, by *definition* the derivative is what the difference quotient approaches as h goes to zero. So let's first simplify the difference quotient:

$$\frac{f(x+h) - f(x)}{h} = \frac{(x+h)^4 - (x)^4}{h}.$$

This is the hardest step – what is $(x+h)^4$? Well, it turns out we can be a little lazy about it; we'll see why in a second. But if you begin writing out what $(x+h)^4$ is, we see

$$(x+h)^4 = (x+h) \times (x+h) \times (x+h) \times (x+h).$$

So I'm going to mysteriously ask: If we multiply all this out, what are the coefficients of x^4, and what are the coefficients of x^3?

To have a term like "x^4" pop up, we need x to be multiplied with itself 4 times:

$$(\underline{x} + h) \times (\underline{x} + h) \times (\underline{x} + h) \times (\underline{x} + h)$$

so multiplying the underlined terms together, we get x^4. But how can we get an x^3 to pop up? Well, there are four ways:

$$(x + \underline{h}) \times (\underline{x} + h) \times (\underline{x} + h) \times (\underline{x} + h)$$

$$(\underline{x}+h)\times(x+\underline{h})\times(\underline{x}+h)\times(\underline{x}+h)$$
$$(\underline{x}+h)\times(\underline{x}+h)\times(x+\underline{h})\times(\underline{x}+h)$$
$$(\underline{x}+h)\times(\underline{x}+h)\times(\underline{x}+h)\times(x+\underline{h})$$

and each of these four ways involves multiplying three xs against one h – so the x^3 term will be of the form $4x^3h$. In other words, we've found that

$$(x+h)^4 = x^4 + 4hx^3 + \text{other stuff}$$

where "other stuff" is the collection of the terms we haven't accounted for yet. For example, we haven't talked about what we'll get if multiply out:

$$(x+\underline{h})\times(\underline{x}+h)\times(\underline{x}+h)\times(x+\underline{h}).$$

But note that the only terms we haven't accounted for are terms with *at least two h* being multiplied together. So whatever "other stuff" is, every term will have an h^2 or an h^3 or an h^4 – some higher power of h – in it.

Upshot: $f(x+h) = x^4 + 4x^3h+$ stuff with higher powers of h. (Here, "higher powers" means the power of h is at least 2.)

We can now return to computing the rest of the difference quotient. First, the numerator:

$$f(x+h) - f(x) = f(x+h) - x^4.$$

So based on our Upshot from above, we see that

$$f(x+h) - f(x) = 4x^3h + \text{stuff with higher powers of } h.$$

Thus

$$\begin{aligned}\frac{f(x+h)-f(x)}{h} &= \frac{4x^3h + \text{stuff with higher powers of } h}{h}\\ &= \frac{4x^3h}{h} + \frac{\text{stuff with higher powers of } h}{h}\\ &= 4x^3 + \frac{\text{stuff with higher powers of } h}{h} \qquad (\text{when } h \neq 0).\end{aligned}$$

Well, the term on the right involves taking a sum of higher powers of h, and dividing by h. The result will be a sum of terms that all involve (multiples of) h, h^2, h^3, et cetera. In particular, as h goes to zero, these terms will all go to zero! To summarize:

$$4x^3 + \frac{\text{stuff with higher powers of } h}{h} \to 4x^3 + 0 \qquad (\text{as } h \text{ approaches } 0).$$

So we see that the derivative is indeed $4x^3$.

Here is a formal proof demonstrating that n doesn't have to equal only 4. The following is a proof[8] for any $n \geq 1$:

Proof of Power Law (Proposition 3.5.1). Let n be a whole number of size at least 1. Let us note that we can compute $(x+h)^n$ as a product of n copies of $x+h$ – and in doing so, we find that

$$(x+h)^n = x^n + nx^{n-1}h + \text{terms with higher powers of } h.$$

Thus,

$$(x+h)^n - x^n = nx^{n-1}h + \text{terms with higher powers of } h.$$

Dividing this expression by h, we see that (when $h \neq 0$):

$$\frac{(x+h)^n - x^n}{h} = nx^{n-1} + \text{terms with } h,\ h^2,\ h^3,\ \ldots.$$

And as h approaches zero, the terms with h, h^2, h^3 all shrink as well; leaving

$$nx^{n-1}.$$

In other words, as h approaches zero, the difference quotient of $f(x) = x^n$ appraoches nx^{n-1}. By definition, we have shown that the derivative of x^n is nx^{n-1}. This is what we wanted to prove. □

[8] A "proof" in mathematics is a demonstration that a fact is true. Most people live life asserting facts willy-nilly without any justification. In mathematics, your professor should be able to justify absolutely everything they tell you. That's why I'm including this bonus section – to justify that the power law is true.