Calculus I: Lecture Notes

Quinn La Fond

May 13, 2025

Contents

Week I: Limits

We make the following notational conventions:

- For an arbitrary function $f(x)$, we will denote by $f^n(x)$ for positive n to be the n-fold product of f. We will denote by $f^{(n)}(x)$ as the nth derivative, and $f^{-1}(x)$ to be the inverse function to

f. If we wish to write something of the form $\frac{1}{f(x)}$, more compactly, we will write $(f(x))^{-1}$.

- If $n > 0$, then $\sin^n x$ is to be the n-fold product of $\sin x$. In other words $\sin^2 x = (\sin x) \cdot (\sin x)$. If $n < -1$, we will never employ the notation $\sin^n x$, and if $n = -1$, then $\sin^{-1} x$ will be the inverse function to sin, often denoted arcsin. The same holds for other trigonometric functions.
- The functions $\csc x$, $\sec x$ and $\cot x$ denote $\frac{1}{\sin(x)}$, $\frac{1}{\cos x}$ and $\frac{1}{\cot x}$ respectively.
- The function $\log x$ will always refer to the base 10 logarithm, while $\ln x$ will be base $e \approx 2.718\ldots$; if we need to use logarithms in base a for some real number a, then will denote them by $\log_a x$.

Furthermore, there are some things in this course we have the ability to prove, some for which we can provide a fake 'almost proof', and some for which we cannot prove in any manner. For statements we can prove, the proof will always be followed by an italicized *proof* and end with a □. For statements which we can provide a fake proof for, we will engage in some discussion motivating the statement with some hand wavy techniques, and then state the result. For statements which we cannot prove in manner, we simply state the result.

1.1 Introduction and Motivation

Suppose you drop a ball off a building that is 500 meters tall, then results from physics tell us that the height of the ball as a function of time is given by:

$$y(t) = -10 \cdot t^2 + 500 \tag{1.1}$$

Note that the above function only makes physical sense on the interval $[0, t_0]$, where t_0 is the moment that the ball hits the ground. Since the height of the ball at t_0 is zero, we can find t_0 by setting (1.1) equal to zero:

$$-10 \cdot t^2 + 500 = 0 \Rightarrow 10t^2 = 500 \Rightarrow t^2 = 50 \Rightarrow t = \pm\sqrt{50} = \pm 5 \cdot \sqrt{2}$$

We also know that we should take the positive square root, as negative time does not make physical sense. It follows that our height function is physically defined on $[0, 5 \cdot \sqrt{5}]$. We now might ask ourselves a variety of different physical questions about our falling ball, such as: what is the average speed of the ball? This is a question we can answer with purely algebraic techniques; indeed we know the ball travels a total of $\Delta y = -500$[1] meters over a span of $\Delta t = 5 \cdot \sqrt{5}$ seconds, so the average speed, v_{avg}, is given by:

$$v_{\text{avg}} = \frac{\Delta y}{\Delta t} = \frac{500}{5 \cdot \sqrt{s}} \approx -70.71\,\text{m/s}$$

However, what if we want to know the speed of the ball when it has traveled 250 meters? Or right before it crashes into the ground? Or at any point along it's trajectory? Answering these questions requires more sophisticated techniques, the techniques of calculus. In fact, the field of calculus was almost entirely motivated by questions of this form.

Our first step in answering such questions is to analyze the average velocity of our ball over a Δt[2] geometrically. Suppose we want to calculate the speed of our ball at $t_1 = 2$, $y(t_1) = 460$, then a good place to start is to consider the average speed of the ball over an interval starting at t_0, say $[2, 5]$. In this case, v_{avg} is given by:

$$v_{\text{avg}} = \frac{\Delta y}{\Delta t} = \frac{y(t_1 + 3) - y(t_1)}{(t_1 + 3) - t_2} = \frac{250 - 460}{3} = -70\,\text{m/s}$$

Now, v_{avg} is the slope of the line which goes through the points:

$$(t_1, y(t_1)) = (2, 460) \qquad \text{and} \qquad (t_1 + 3, y(t_1 + 3)) = (5, 250)$$

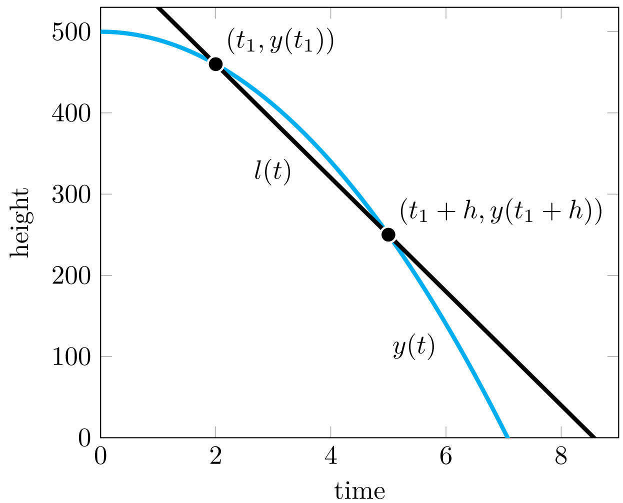
[1] Here Δy is negative as $\Delta y = y_f - y_i = 0 - 500$

[2] In generality Δ means change in some quantity. In this case it the final time minus the initial time.

The function corresponding to this line[3] is given by:

$$l(t) = v_{\text{avg}} \cdot (t-2) + 460$$

We can then draw the following graph:



The black line is $l(t)$, the blue line is $y(t)$, and we have marked two points on the graph, $(t_1, y(t_1))$ and $(t_1 + h, y(t_1 + h))$, where $h = 3$. The key insight is that if we could somehow take the average speed over the interval $[t_1, t_1]$, we would obtain the speed of the ball at $(t_1, y(t_1))$ because that interval consists of only a single point. We cannot do this naively though, as our formula for average velocity, i.e. the slope of the line passing through the end points of the interval would yield:

$$\frac{\Delta y}{\Delta t} = \frac{y(t_1) - y(t_1)}{t_1 - t_1} = \frac{0}{0}$$

which doesn't make mathematical or physical sense! So, how can we rectify the situation? The next key insight is that if we allow h to vary instead of being fixed, then as h gets closer and closer to zero, since the interval becomes smaller and smaller, we get average velocities which are closer and closer to the velocity at t_0. Allowing h to vary makes v_{avg} a function of h given by:

$$v_{\text{avg}}(h) = \frac{\Delta y}{\Delta t} = \frac{y(t_1 + h) - y(t_1)}{t_1 + h - t_1} = \frac{-10(t_1 + h)^2 + 500 - (-10t_1^2 + 500)}{h} = \frac{-10h^2 - 20h \cdot t_1}{h}$$

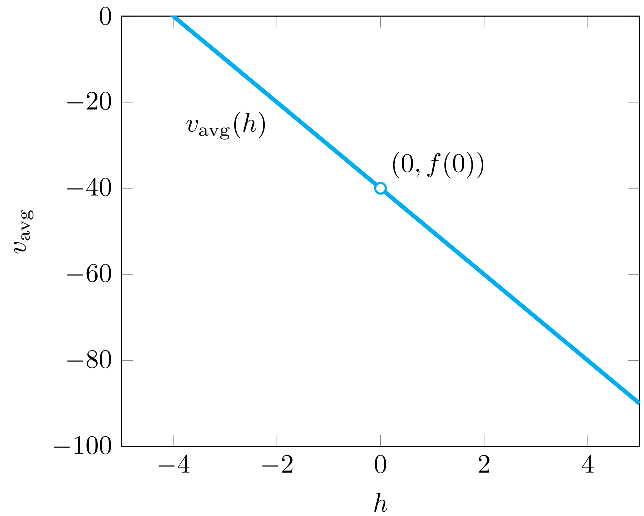
This is a *rational function*, and as such is not defined at $h = 0$, however, for all $h \neq 0$, we have that:

$$v_{\text{avg}}(h) = -10h - 20 \cdot t_1 \tag{1.2}$$

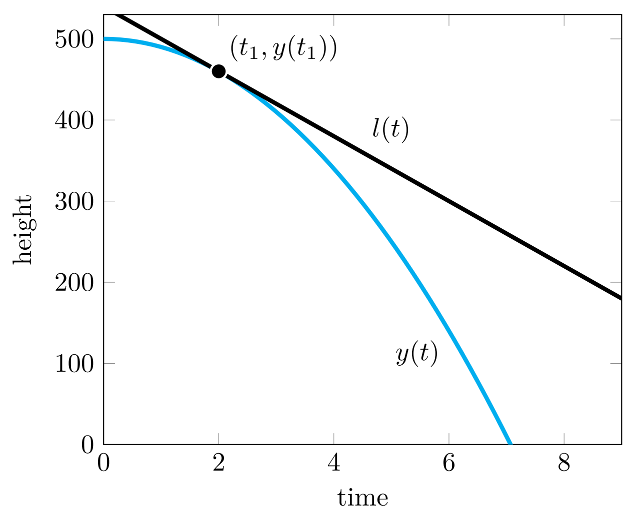
because each term has at least one power of h, and we can divide by h when h is nonzero. In particular,

[3] i.e. the function whose graph is this line.

the graph of this function is the graph of the function $f(h) = -10h - 20 \cdot t_1$ with a hole at $h = 0$:



So while we can't actually plug $h = 0$ into our v_{avg} to get the speed at t_1, it is clear from the above graph, that as h gets closer to 0, $v_{\text{avg}}(h)$ approaches the numerical $f(0) = -20 \cdot t_1$, which in this case is -40 as $t_1 = 2$. It therefore makes intuitive sense to say that the speed of the ball at t_1 is $-60\,\text{m/s}$. Moreover, if we replace $l(t)$ in our original picture with the line going through the point $(-2, 460)$, at a slope $m = -60$, we have the following:



So the velocity of the ball at $t_1 = 2$ is also the slope of the line *tangent*[4] to the graph of $y(t)$ at $(t_1, y(t_1))$.

The entire process outlined above is called 'taking the derivative at t_1'. In particular, the process of seeing what the value of $v_{\text{avg}}(h)$ is as h approaches 0 (even though $v_{\text{avg}}(h)$ is not defined at $h = 0$!) is called 'taking the limit of $v_{\text{avg}}(h)$ as h goes to zero'. We employ the following notation for this:

$$\lim_{h \to 0} v_{\text{avg}}(h)$$

In particular, if we let t_1 vary, we get a new function defined by:

$$y'(t) = \lim_{h \to 0} \frac{y(t+h) - y(t)}{h}$$

[4]By which we mean only glances the graph of the function at $(2, 440)$, instead of intersecting it.

By our earlier work, this is the same as:

$$y'(t) = \lim_{h \to 0} \frac{-10h^2 - 20h \cdot t}{h}$$

For all nonzero h this is equal to the equation (1.2), so the limit as h approaches zero is precisely $-20 \cdot t$. It follows that we have a function:

$$y'(t) = -20 \cdot t$$

This is called the derivative of $y(t)$, and for every t_1 in the interval $[0, 5 \cdot \sqrt{2}]$, when we plug in t_1 to $y'(t)$, we get the velocity of the ball at the time t_1, or equivalently the velocity of the ball at a height of $y(t_1)$ meters.

The rest of the course, and much of calculus in general, is about studying the properties of derivatives for various functions, but even when we delve deeper into abstraction, and away from the world of physics, we should keep the above picture of a ball falling off a building in mind.

1.2 Definitions and Examples

In the previous section, we motivated the idea of a derivative by examining a physical problem. However, the process for finding the speed of a ball as it falls from a building relied on the notion of 'taking a limit' of a function, and in fact the concept of a derivative relies heavily on this idea. Due to this we spend the next few sections, discussing limits, and their properties.

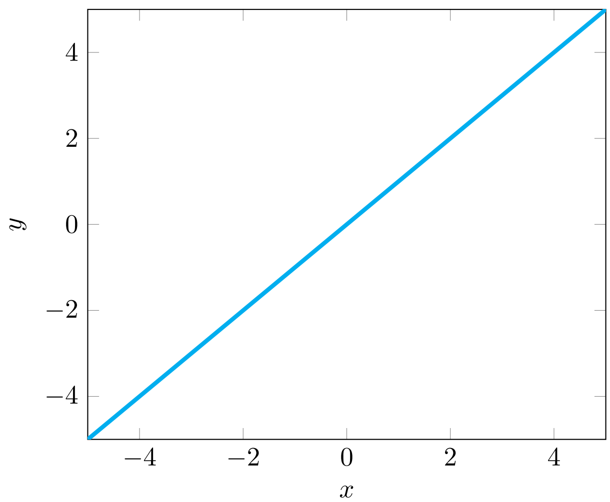
Limits essentially come in two flavors, we have limits as 'x goes to positive or negative infinity', and limits as 'x approaches a', where a is some finite real number. The former is a way of characterizing the long term behavior of a function $f(x)$, and the later analyzes the behavior of f near a point a, even if $f(a)$ is undefined. We begin with limits of the form x goes to positive or negative infinity, and employ the notation:

$$\lim_{x \to \infty} f(x) \qquad \text{and} \qquad \lim_{x \to -\infty} f(x)$$

Now limits of the above form can be a real number, can be 'positive or negative infinity', or they can be non existent. When limits above the form are a real number say a, this means that as x gets bigger and bigger, (or more and more negative), that the value of $f(x)$ gets closer and closer to a. In other words, in this situation we have that as x approaches positive (or negative) infinity, $f(x)$ approaches a. The next option is that $f(x)$ 'blows up' as x approaches positive or negative infinity, by which we mean that x gets bigger and bigger (or more and more negative) the value of $f(x)$ continues to grow in either the positive or negative direction. In this case, we have the limit as x approaches positive or negative infinity is equal to positive or negative infinity, depending on which direction the function grows. The final option is that the limit may fail to exist, in which case we simply write DNE. This can happen if the long term behavior of the function is oscillatory like if $f(x) = \sin x$; in this case the function neither grows without bound, nor does it approach some constant, it just oscillates between 1 and -1 periodically, hence there is no value that $f(x)$ approaches as x approaches positive or negative infinity.

The quickest way of dealing with limits of these form is by using 'big number logic'. This is best taught via example:

Example 1.1. Let $f(x) = x$, then the graph of $f(x)$ is:



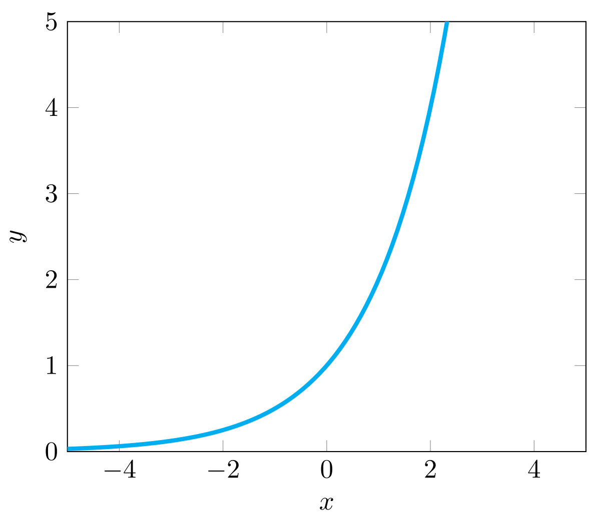
Now as we plug larger and larger and numbers into $f(x)$, we just return that same number since $f(x) = x$. In particular as x gets bigger $f(x)$ bigger, so the long term behavior of $f(x)$ is to get bigger and bigger. When this occurs, we write:

$$\lim_{x \to \infty} f(x) = \infty$$

because $f(x)$ will just keep growing as x grows. Similarly, when x gets more and more negative, $f(x)$ gets more and more negative, hence:

$$\lim_{x \to -\infty} f(x) = -\infty$$

Example 1.2. Let $f(x) = 2^x$, we can see from the graph of this function:



that the limit as x approaches positive infinity is infinity, and that the limit as x approaches negative infinity is 0. But how can we determine this with big number logic? The idea is that if x is getting bigger, i.e. as x approaches positive infinity, then 2^x also just gets bigger, as we are just taking larger and larger powers of two. From this we conclude that:

$$\lim_{x \to \infty} 2^x = \infty$$

However, if x is negative, then we are dividing 1 by larger and larger powers of 2. In particular, we have the following infinite sequence when x is a negative integer:

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots$$

So as we plug in more and more negative values for x, 2^x gets closer and closer to zero, because we get fractions with larger and larger denominators. It follows that:

$$\lim_{x \to \infty} 2^x = 0$$

Example 1.3. Let:

$$f(x) = \frac{5x^3 + 2x^2 - x + 1}{x^2 + 3x - 2}$$

We want to determine the limit as $x \to \infty$ and $x \to -\infty$. Unlike the previous two cases, this function is not easily graphed by hand, so we have to rely solely on big number logic. Let us first determine the limit as $x \to \infty$; the point is that as x gets very very large ,the terms which contribute most to the quotient:

$$\frac{5x^3 + 2x^2 - x + 1}{x^2 + 3x - 2}$$

are only the highest order terms in numerator and denominator. In other words, if x is very very large, then $5x^3 + 2x^2 - x + 1$ is very close to $5x^3$, because x^3 will be so much larger than $2x^2 - x + 1$. The same holds for the denominator, hence:

$$\begin{aligned} \lim_{x \to \infty} \frac{5x^3 + 2x^2 - x + 1}{x^2 + 3x - 2} &= \lim_{x \to \infty} \frac{5x^3}{x^2} \\ &= \lim_{x \to \infty} 5x \\ &= \infty \end{aligned}$$

Similarly, the same logic demonstrates that:

$$\begin{aligned} \lim_{x \to -\infty} \frac{5x^3 + 2x^2 - x + 1}{x^2 + 3x - 2} &= \lim_{x \to \infty} \frac{5x^3}{x^2} \\ &= \lim_{x \to -\infty} 5x \\ &= -\infty \end{aligned}$$

Note that if we change the denominator to be a cubic:

$$f(x) = \frac{5x^3 + 2x^2 - x + 1}{x^3 + 3x - 2}$$

then:

$$\begin{aligned} \lim_{x \to \infty} \frac{5x^3 + 2x^2 - x + 1}{x^3 + 3x - 2} &= \lim_{x \to \infty} \frac{5x^3}{x^3} \\ &= \lim_{x \to \infty} 5 \\ &= 5 \end{aligned}$$

and:

$$\begin{aligned} \lim_{x \to -\infty} \frac{5x^3 + 2x^2 - x + 1}{x^3 + 3x - 2} &= \lim_{x \to -\infty} \frac{5x^3}{x^3} \\ &= \lim_{x \to -\infty} 5 \\ &= 5 \end{aligned}$$

If we make the denominator a quartic:

$$f(x) = \frac{5x^3 + 2x^2 - x + 1}{x^4 + 3x - 2}$$

then:

$$\begin{aligned}
\lim_{x\to\infty} \frac{5x^3 + 2x^2 - x + 1}{x^3 + 3x - 2} &= \lim_{x\to\infty} \frac{5x^3}{x^4} \\
&= \lim_{x\to\infty} \frac{5}{x} \\
&= 0
\end{aligned}$$

and:

$$\begin{aligned}
\lim_{x\to-\infty} \frac{5x^3 + 2x^2 - x + 1}{x^3 + 3x - 2} &= \lim_{x\to-\infty} \frac{5x^3}{x^4} \\
&= \lim_{x\to-\infty} \frac{5}{x} \\
&= 0
\end{aligned}$$

In particular, big number logic gives us the following result:

Theorem 1.1. *If $p(x)$ and $q(x)$ are polynomials such that:*

$$p(x) = a_n x^n + \cdots + a_1 x + a_0 \qquad \text{and} \qquad q(x) = b_m x^m + \cdots + b_1 x + b_0$$

for some positive integers m and n, and real numbers $a_0, \ldots, a_n, b_0, \ldots b_n$, then:

$$\lim_{x\to\infty} \frac{p(x)}{q(x)} = \lim_{x\to\infty} \frac{a_n x^n}{b_m x^m} \qquad \text{and} \qquad \lim_{x\to-\infty} \frac{p(x)}{q(x)} = \lim_{x\to-\infty} \frac{a_n x^n}{b_m x^m}$$

Example 1.4. Earlier we noted that $\lim_{x\to\infty} \sin x$ does not exist due to its oscillatory behavior. In this example, we examine a similar function:

$$f(x) = \frac{\sin x}{e^x}$$

Using big number logic, we see that as x gets very large, we are dividing numbers between -1 and 1, i.e. $\sin x$, by an extremely large number e^x. It follows that even though the function is oscillating, it is getting closer and closer to zero as x approaches infinity, so:

$$\lim_{x\to\infty} \frac{\sin x}{e^x} = 0$$

However, as x gets more and more negative, we are dividing a number between -1 and 1, i.e. $\sin x$, by a number getting closer and closer to zero since $\lim_{x\to-\infty} e^x = 0$. It follows that $\sin x/e^x$ is oscillating between extremely large negative numbers and extremely large positive numbers, hence no limit exists, as it is not growing in a consistent direction. Therefore:

$$\lim_{x\to-\infty} \frac{\sin x}{e^x} \text{ does not exist}$$

We now begin our handling of limits as 'x approaches a' for some real number a. Instead of writing 'limit of $f(x)$ as x approaches a' we employ the notation:

$$\lim_{x\to a} f(x)$$

Just as the infinite limits, the 'result' of the above expression has three possibilities, all of which tell us something about the behavior of $f(x)$ near $x = a$. The first possibility is that:

$$\lim_{x \to a} f(x) = L$$

for some real number L; what this means is that as x approaches, or gets closer and closer to a, the values $f(x)$ get closer and closer to a. Now note that that x could approach a from the left or the right of a, so for the limit to be equal to L, $f(x)$ has to approach L in both directions; we will delve more into this later. An example of this case is our $v_{\text{avg}}(h)$ function from Section 1.1; as h approached 0 the value of $v_{\text{avg}}(h)$ approached the speed at which the ball was traveling at $t_0 = 2$. In particular:

$$\lim_{h \to 0} v_{\text{avg}}(h) = -40$$

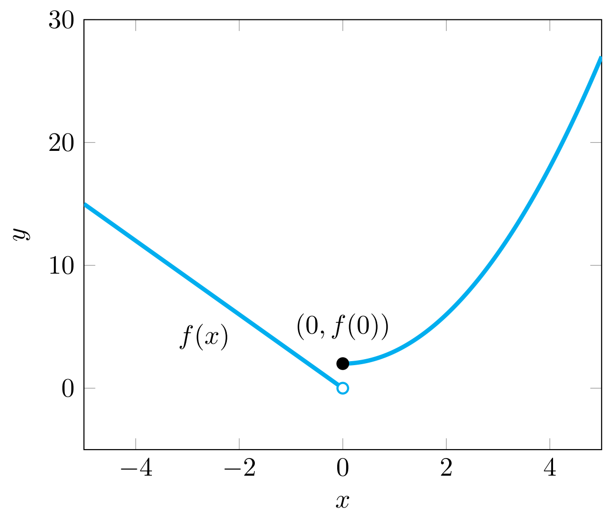
The next situation is that:

$$\lim_{x \to a} f(x) = \pm\infty$$

by which we mean that as x approaches a, $f(x)$ grows without bound in the positive or negative direction. In other words $f(x)$ gets larger and larger, or more and more negative as x approaches a. Finally, we can have that the limit of $f(x)$ as x approaches a fails to exist. This is most commonly found in the following situation, let:

$$f(x) = \begin{cases} x^2 + 2 & \text{for } x \geq 0 \\ -3x & \text{for } x < 0 \end{cases} \tag{1.3}$$

The above notation means that for $x < 0$, $f(x) = -3x$ and for $x \geq 0$ $f(x) = x^2 + 2$. The graph of this function is given by:



Now what should the limit of $f(x)$ be as x goes to zero? The problem is that if $x > 0$, then as x gets closer and closer to zero, $f(x)$ gets closer and closer to 2, but if $x < 0$ then as x get closer and closer to zero, $f(x)$ gets closer and closer to 0. We clearly don't have that $f(x)$ is growing in a consistent direction, and from our earlier discussion, $f(x)$ can't approach a consistent value L, so the limit as x approaches $f(x)$ does not fall into either of the previously discussed categories. In this case, we thus say the limit as x approaches a of $f(x)$ does not exist. Notation we say that:

$$\lim_{x \to a} f(x) \text{ does not exist}$$

Before delving into examples, we briefly formalize our analysis of the limit of $f(x)$ as defined in (1.3).

Definition 1.1. Let $f(x)$ be a function, and a a real number. We define the **limit as x approaches a from the left** as a limit of $f(x)$ where we only consider values of $x < a$. We denote this by:

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

In other words, we only care if $f(x)$ approaches L, grows in a positive or negative direction, or does not exist while analyzing values of x which are less than a. Similarly we define the **limit as x approaches a from the right** as a limit of $f(x)$ where we only consider values of $x > a$. We denote this by:

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

In particular, if $f(x)$ is as defined in (1.3) we have that:

$$\lim_{x \to a^-} f(x) = 0 \neq 2 = \lim_{x \to a^+} f(x)$$

We have the following result:

Theorem 1.2. *Let $f(x)$ be a function, and a a real number. Then the following are true:*

a) If $\lim_{x \to a} f(x) = L$, ∞, or $-\infty$, for some real number L, then $\lim_{x \to a^-} f(x) = \lim_{x \to a^+} f(x) =L$, ∞ or $-\infty$ respectively.

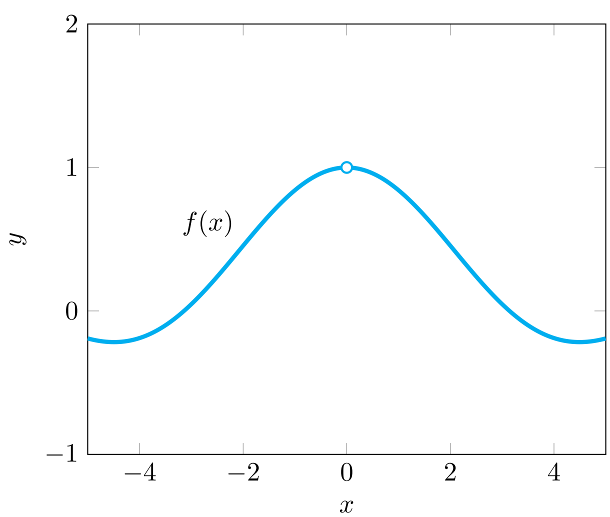
b) If $\lim_{x \to a^-} = \lim_{x \to a^+} = L$, ∞ or $-\infty$, then $\lim_{x \to a} f(x) = L$, ∞, or $-\infty$ respectively.

We now look at some examples:

Example 1.5. Let:

$$f(x) = \frac{\sin x}{x}$$

then the graph of $f(x)$ is given by:



From the graph of the function, it is easy to see that as x approaches 0 from the left, $f(x)$ approaches 1, and as x approaches 0 from the right, $f(x)$ approaches 1. We thus have that:

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

This was easy since we could see the graph, but graphing this function by hand is difficult (I know I wouldn't be able to do it!); on the first homework you will calculate this limit formally with appealing to the graph of the function.

1.3 Continuity

Note even if $f(a) = L$ it could be the case that $\lim_{x\to a} f(x) \neq L$. Indeed consider the next example:

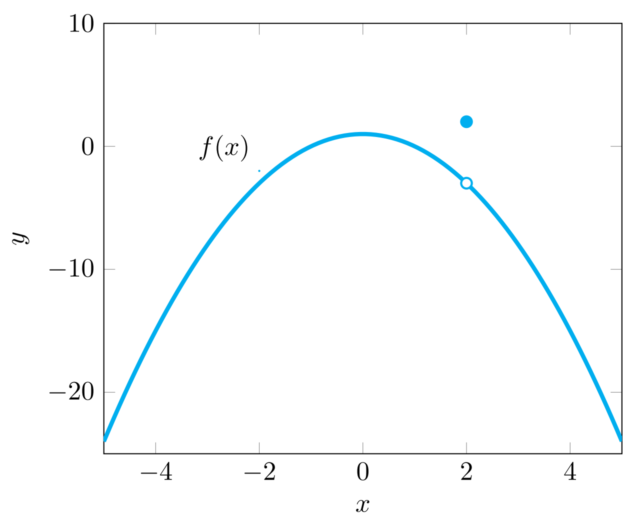
Example 1.6. Let:

$$f(x) = \begin{cases} -x^2+1 & \text{if } x \neq 2 \\ 2 & \text{if } x = 2 \end{cases}$$

Let us analyze the limit as x approaches 2 without appealing to a graphical argument. Since x^2+1 is a continuous function, i.e. we can draw it's graph without lifting up our pencil, we have that as x gets closer and closer to 2, x^2+1 gets closer and closer to 5. One can see this with the following chart:

x	$-x^2+1$
1.9	-2.61
1.99	-2.96
2.01	-3.04
2.1	-3.41

It follows that $\lim_{x\to 2} f(x) = -3$, however from the definition of $f(x)$, we have that $f(2) = 2$. Looking at the graph of this function, we see that the fact that $\lim_{x\to 2} f(x) \neq f(2)$ reflects the fact that $f(x)$ is not a continuous function:



With this example in mind, we give a different definition of a function being continuous:

Definition 1.2. Let $f(x)$ be a function, then $f(x)$ is **continuous at a** if:

$$\lim_{x\to a} f(x) = f(a)$$

A function $f(x)$ is **continuous on it's domain** if for every real number a in the domain of $f(x)$, $f(x)$ is continuous at a. A function is **continuous** if it is continuous for every real number. A **discontinuity of f(x)** is a real number a such that $f(x)$ is not continuous at a, or $f(x)$ is not defined at a. A discontinuity a of $f(x)$ is a **removable discontinuity**, if:

$$\lim_{x\to a} f(x) = L$$

for some real number L. A discontinuity a of $f(x)$ is a **jump discontinuity** if:

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

for some real numbers L^- and L^+. A discontinuity a of $f(x)$ is an **infinite discontinuity** if:

$$\lim_{x\to a^+} f(x) = \pm\infty \qquad \text{or} \qquad \lim_{x\to a^-} f(x) = \pm\infty$$

This definition, while more verbose and complicated than the 'drawing a graph without lifting up a pen' definition, mathematically captures the spirit of the concept of continuity, and is therefore the 'correct' definition for this concept. Trigonometric functions, exponential functions, logarithmic functions, radical functions[5] and rational functions are all continuous on their domains; that is they are continuous everywhere they are defined. Polynomials, exponential functions, $\sin x$ and $\cos x$ are examples of continuous functions, as they are defined everywhere.

Example 1.7. Let $f(x)$ be the function from Example 1.6, then $f(x)$ is continuous everywhere but $x = 2$. Indeed, at $x = 2$ we have that $\lim_{x\to 2} f(x) = -3$ but $f(2) = 2$. It follows that 2 is a discontinuity because $f(x)$ is not continuous at 2. In particular, 2 is a removable discontinuity, because the limit as x approaches 2 exists and is finite. This example demonstrates why it is called a removable discontinuity, because we can alter the value of the function at one point to make $f(x)$ continuous there.

Example 1.8. Let:

$$f(x) = \frac{x^2 - 9}{x + 3}$$

then $f(x)$ is not defined at $x = -3$ as we will divide by zero. However, for all values $x \neq -3$, we have that:

$$\frac{x^2 - 9}{x + 3} = \frac{(x + 3)(x - 3)}{x + 3} = x - 3$$

It follows that $\lim_{x\to -3} = -6$, so $f(x)$ is not continuos at $x = -3$, because $f(x)$ is not defined at $x = -3$ but it's limit exists, so -3 is a removable discontinuity. Indeed, if we define:

$$g(x) = \begin{cases} \frac{x^2-9}{x+3} & \text{if } x \neq -3 \\ -6 & \text{if } x = -3 \end{cases}$$

then $g(-3) = -6$, and $\lim_{x\to -3} g(x) = -6$ so $g(x)$ is continuous. We have in a sense removed the discontinuity with $g(x)$.

Example 1.9. Consider the function $f(x)$ as defined in Equation 1.3. Then the limit as x approaches 0 of $f(x)$ does not exist, so $f(x)$ is not continuous at $x = 0$. It follows that 0 is a discontinuity point, and it is a jump discontinuity because $\lim_{x\to 0^+} f(x) = 2$ and $\lim_{x\to 0^-} f(x) = 0$. The graph of $f(x)$ demonstrates why we call such a discontinuity a jump discontinuity, because $f(x)$ 'jumps' from one value to the next at $x = 0$.

Example 1.10. Let:

$$f(x) = \frac{1}{x}$$

then $f(0)$ is undefined as we would be dividing by zero. As we approach 0 from the left, we are dividing by negative numbers closer and closer to zero, hence $f(x)$ is approaching negative infinity. As we approach 0 from the right, we are dividing by smaller and smaller positive numbers, so $f(x)$ is 'blowing up' and approaching positive infinity. It follows that:

$$\lim_{x\to 0^+} f(x) = \infty \neq -\infty = \lim_{x\to 0^-} f(x)$$

So we have that the limit as x approaches zero does not exist, that $x = 0$ is a discontinuity point, and in particular it is an infinite discontinuity. If instead:

$$f(x) = \frac{1}{x}^2$$

then we have that the limit as x approaches 0 is ∞ because the left and right handed limits agree, however 0 is still an infinite discontinuity of $f(x)$.

[5] i.e. any function of the form x^a where a is not a whole number

We end the section with the following result on limits, known as the limit rules, and then use them to compute some examples.

Theorem 1.3. *Let a be a real number, and $f(x)$ and $g(x)$ defined for all $x \neq a$ on some open interval containing a. Moreover, suppose that*

$$\lim_{x\to a} f(x) = L \qquad \textit{and} \qquad \lim_{x\to a} g(x) = M$$

for some real numbers L and M, then the following results hold:

a) $\lim\limits_{x\to a}(f(x)+g(x)) = \lim\limits_{x\to a} f(x) + \lim\limits_{x\to a} g(x) = L + M$.

b) $\lim\limits_{x\to a}(f(x)-g(x)) = \lim\limits_{x\to a} f(x) - \lim\limits_{x\to a} g(x) = L - M$.

c) $\lim\limits_{x\to a}(f(x)\cdot g(x)) = \lim\limits_{x\to a} f(x) \cdot \lim\limits_{x\to a} g(x) = L \cdot M$

d) If $M \neq 0$, then:

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim\limits_{x\to a} f(x)}{\lim\limits_{x\to a} g(x)} = \frac{L}{M}$$

e) If $f(x)$ is continuous at M, then

$$\lim_{x\to a} f(g(x)) = f\left(\lim_{x\to a} g(x)\right) = f(M)$$

The above rules, and your challenge homework imply that polynomials are continuous for all real numbers. We now show that rational functions are continuous on their domain; showing that radical functions are continuous on their domain is a fact we take for granted as it is harder to show for *all* real numbers.

Example 1.11. Let $f(x)$ be a rational function, then $f(x)$ is of the form:

$$\frac{p(x)}{q(x)}$$

for some polynomials $p(x)$ and $q(x)$. The domain of $f(x)$ is all real numbers such that $q(x) \neq 0$, so for any a such that $q(a) \neq 0$, we have that:

$$f(a) = \frac{p(a)}{q(a)}$$

Since polynomials are continuous, we have that by d) of the limit rules:

$$\lim_{x\to a} f(x) = \frac{\lim\limits_{x\to a} p(x)}{\lim\limits_{x\to a} q(x)} = \frac{p(a)}{q(a)}$$

so $f(x)$ is continuous on it's domain because it is continuous at every real number for which it is defined.

We claimed earlier that $\sin x$ and $\cos x$ were continuous functions by appealing to their graphs. However, every graph we draw is really only over some interval (a, b), so we have only shown that $\sin x$ and $\cos x$ are continuous over some some small interval, usually including 0. In the following example, we show that $\sin x$ is continuous at all real numbers:

Example 1.12. Let a be a real number, then we need to show that $\lim\limits_{x\to a} \sin x = \sin a$. Note that $\sin x = \sin((x-a)+a)$ because $x = x - a + a$. Using the trigonometric identity:

$$\sin(\theta + \gamma) = \sin\theta\cos\gamma + \cos\theta\cos\gamma$$

we find that:

$$\sin x = \sin(x-a)\cos a + \cos(x-a)\sin a$$

Using the first limit rule, we find that:

$$\lim_{x\to a} \sin x = \lim_{x\to a}(\sin(x-a)\cos a) + \lim_{x\to a}(\cos(x-a)\sin a)$$

We can view $\sin a$ and $\cos a$ as constant functions, hence using c) of Theorem 1.3, and the fact that constant functions are continuous, we have:

$$\lim_{x\to a} \sin x = \cos a \cdot \lim_{x\to a} \sin(x-a) + \sin a \cdot \lim_{x\to a} \cos(x-a)$$

Since $\lim_{x\to a} x-a$ is equal to zero as $x-a$ is a continuous function, and $\sin x$ and $\cos x$ are continuous at zero[6], we have that by e):

$$\begin{aligned}\lim_{x\to a} \sin x &= \cos a \cdot \sin\left(\lim_{x\to a} x-a\right) + \sin a \cdot \cos\left(\lim_{x\to a} x-a\right) \\ &= \cos a \cdot \sin 0 + \sin a \cdot \cos 0 \\ &= \sin a\end{aligned}$$

meaning that $\sin x$ is continuous!

Using the limit laws, and properties of continuous functions, we can calculate many limits, but what about when the limit laws don't apply? For example, so suppose that $f(x)$ and $g(x)$ are functions, satisfying $\lim_{x\to a} f(x) = 0$, and $\lim_{x\to a} g(x) = 0$? Then if we naively try to apply the limit laws to their quotient we get:

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

which doesn't make any sense. We have already seen how to deal with problem in certain cases such as Example 1.8, but we now provide an example of a more complicated situation.

Example 1.13. Let:

$$f(x) = \frac{x-2}{\sqrt{8-x^2}-2}$$

Note that there are two constraints on the domain of this function, namely that $8-x^2$ is greater than or equal to zero, and that $\sqrt{8-x^2}$ does not equal to 2. It follows that the domain of this function is given by[7]:

$$\left[-2\sqrt{2}, -2\right) \bigcup \left(-2, 2\right) \bigcup \left(2, 2\sqrt{2}\right]$$

We want to find the limit of $f(x)$ as x approaches 2. Both the top function and the bottom function are continuous at $x=2$, but if we naively apply the limit laws, then we end up with $0/0$, which as we mentioned earlier is no good. Instead, we should algebraically manipulate the equation by noticing we can 'rationalize the denominator'. Recall the difference of squares formula:

$$(a-b)(a+b) = a^2 - b^2$$

If we set $a = \sqrt{8-x^2}$, and $b = -2$, then we have that:

$$(\sqrt{8-x^2}-2)(\sqrt{8-x^2}+2) = 8 - x^2 - 4 = 4 - x^2$$

[6] You can draw their graph on the interval $(-\pi, \pi)$ to see this!

[7] We do this by first finding the interval on which $8-x^2 \geq 0$, and then removing the solution to $\sqrt{8-x^2} = 2$ from said interval.

It follows that for all x:

$$\begin{aligned}
f(x) =& \frac{x-2}{\sqrt{8-x^2}-2} \cdot \frac{\sqrt{8-x^2}+2}{\sqrt{8-x^2}+2} \\
=& \frac{(x-2)(\sqrt{8-x^2}+2)}{4-x^2} \\
=& \frac{(x-2)(\sqrt{8-x^2}+2)}{(2-x)(2+x)} \\
=& \frac{(x-2)(\sqrt{8-x^2}+2)}{-(x-2)(2+x)}
\end{aligned}$$

where in the third step we have applied the difference of squares formula again, and in the final step we pulled out a negative 2. For all $x \neq 2$ we can set $(x-2)/(x-2)$ equal to one, so this simplifies to:

$$f(x) = \frac{(\sqrt{8-x^2}+2)}{-(2+x)}$$

Now we can apply the limit laws as both the top and the bottom functions have non zero limits x approaches 2. Therefore:

$$\begin{aligned}
\lim_{x\to 2} f(x) =& \frac{\lim_{x\to a}\left(\sqrt{8-x^2}+2\right)}{\lim_{x\to a} -(2+x)} \\
=& \frac{\sqrt{8-4}+2}{-(2+2)} \\
=& -1
\end{aligned}$$

This process for solving limits is called rationalizing the denominator.

1.4 The Squeeze Theorem

In this section we go over the Squeeze Theorem; this is conceptually a vital tool to the study of limits, and many of the limits we encounter naturally in our study of calculus will rely on it. Usually, the arguments surrounding the use of the squeeze theorem are tricky, that is why problem II on your challenge homework has you work through some examples on your own (though some of these examples are found in your textbook).

The theorem is as follows:

Theorem 1.4. *Let a be a real number, and $f(x)$ a function defined for all $x \neq a$ on an interval containing a. If $g(x)$ and $h(x)$ are functions defined for all $x \neq a$ on an interval containing a satisfying:*

$$\lim_{x\to a} g(x) = L = \lim -x \to ah(x)$$

and for all $x \neq a$:

$$g(x) \leq f(x) \leq h(x)$$

then $\lim_{x\to a} f = L$ as well. Similarly if:

$$\lim_{x\to a^\pm} g(x) = L = \lim -x \to a^\pm h(x)$$

and for all $x < a$ (for left hand limits) or $x > a$ (for right hand limits)

$$g(x) \leq f(x) \leq h(x)$$

then $\lim_{x\to a^\pm} f = L$ as well.

What exactly is this theorem saying? Well it is saying that if the value of $f(x)$ at every point lies between $g(x)$ and $h(x)$ at that point, then as x approaches a, $f(x)$ must approach L because so do $g(x)$ and $h(x)$! Say for example $L = 3$, if $f(x)$ is constantly greater than something approaching 3, and less than something approaching 3, then f must also approach 3, as there are no other between 3 and 3. This should feel intuitively obvious once you start playing around with the idea. I will show you one harder example, and leave the rest of the exploration up to you on the homework set[8]

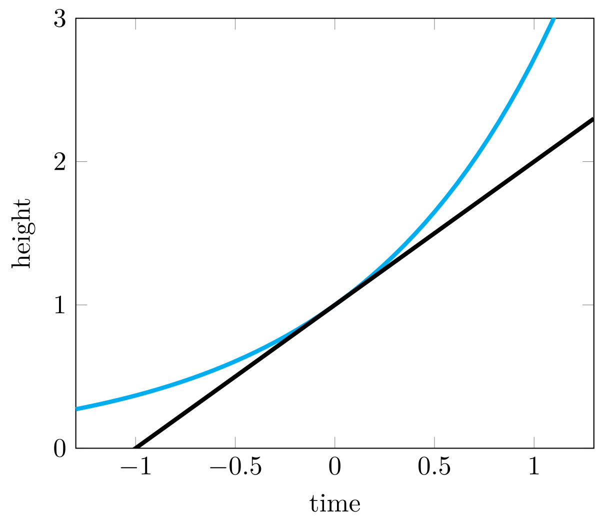
Example 1.14. The following function will be of great importance to us in the future:

$$f(x) = \frac{e^x - 1}{x}$$

We want to find the limit of the above function as x approaches 0. If we plug in 0 then we get 0/0 so the limit laws can't help us here, and unfortunately there is no clever algebraic trick we can use to make the limit tractable as in Example 1.13. We will have to use the squeeze theorem. We will have to use the fact:

$$1 + x < e^x$$

for $-1 < x < 1$ which can be seen from the following graph:



If we replace x with $-x$ we get that:

$$1 - x < e^{-x}$$

Multiplying both sides by e^x, and $1/(1-x)$ we obtain the following inequality:

$$e^x < \frac{1}{1-x}$$

We thus have that:

$$1 + x < e^x < \frac{1}{1-x}$$

subtracting 1 from both sides we get that:[9]

$$x < e^x - 1 < \frac{x}{1-x}$$

[8]Don't worry, I won't ever ask you to use the squeeze theorem on an exam or a daily warm up.

[9]The last part of the inequality comes from the fact that $\frac{1}{1-x} - 1 = \frac{1}{1-x} - \frac{1-x}{1-x} = \frac{x}{1-x}$

Now, we want to divide by x, to get the desired inequality, but we have to be careful. When we divide by positive x nothing changes, so for $0 < x < 1$ we have:

$$1 < \frac{e^x - 1}{x} < \frac{1}{1-x}$$

However, when we divide by x for $x < 0$ the inequality changes direction because dividing by a negative number changes the sign.[10] It follows that for $-1 < x < 0$ we have that:

$$\frac{1}{1-x} < \frac{e^x - 1}{x} < 1$$

Now we have that both $1/(1-x)$ and 1 are continuous on their domain, hence their limits exits at 0, and are both equal to 1. It follows by the squeeze theorem that:

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

hence

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

Week II: Intro to Derivatives

2.1 Definition and Examples

In Section 1.1 we calculated the instantaneous velocity of a ball as it fell from a building. More precisely, we had a function:

$$y(t) = -10t^2 + 500$$

which gave us the height at which the ball was at any time $t \geq 0$. We argued that the instantaneous velocity at $t_0 = 2$ *should be* the average velocity of the ball over the interval $[2, 2]$. The problem is that the average velocity over this interval is:

$$\frac{\Delta y}{\Delta t} = \frac{y(2) - y(2)}{2-2} = \frac{0}{0}$$

which doesn't make sense. Our fix was to define a function $v_{\text{avg}}(h)$ which gives us the average velocity of any interval of the form $[2, 2+h]$, and then argue that as h got closer to zero, $v_{\text{avg}}(h)$ approached a finite value which *should be* the instantaneous velocity. With our newfound language of limits, we can phrase this as follows: the instantaneous velocity of the ball at $t_0 = 2$, denoted $v(2)$ is given by:

$$v(2) = \lim_{h\to 0} \frac{y(2+h) - y(2)}{h}$$

With this we can define a velocity function by:

$$\begin{aligned} v(t) &= \lim_{h\to 0} \frac{y(t+h) - y(t)}{h} \\ &= \lim_{h\to 0} \frac{-10(t+h)^2 + 500 + 10t^2 - 500}{h} \\ &= \lim_{h\to 0} \frac{-10t^2 - 20th - 10h^2 + 10t^2}{h} \\ &= \lim_{h\to 0} \frac{-20th - 10h^2}{h} \\ &= \lim_{h\to 0} -20t - 10h^2 \\ &= -20t \end{aligned}$$

With these results in mind, we employ the following definition:

[10] Think about what happens if you multiply $3 < 5$ by -1, do you get $-3 < -5$? or $-5 < -3$?

Definition 2.1. Let $f(x)$ be a function, then f is **differentiable at a** if the limit:

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

is equal to a real number we denote by $f'(a)$. If this limit does not exist, or is infinite, we say that f is **not differentiable at a**. We call $f'(a)$ the **derivative of f at a**, and define a function:

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

which we call **the derivative of f.** The domain of this function consists of all real numbers where f is differentiable. The **second derivative**, is the derivative of the derivative, and is denoted f''. It comes from taking the derivative twice. We can do this any amount of times actually, and denote the **nth derivative** by $f^{(n)}$.

Note that for f to be differentiable at a it must be defined at a. Furthermore, we remark that some people employ the Leibniz notation:

$$f' = \frac{df}{dx} \qquad \text{and} \qquad f'(a) = \frac{df}{dx}|_a = \frac{df}{dx}(a)$$

for the derivative of f and the derivative of f at a respectively. This notation comes from the idea of of the derivative being an 'infinitesimal average rate of change'. In particular, the average rate of change over some interval $[a,b]$ is given by:

$$\frac{\Delta f}{\Delta x} = \frac{f(b)-f(a)}{b-a}$$

and so our limit definition of a derivative, is like looking at infinitely small changes in f, called df, divided by infinitely small changes in x called dx. This is a fine way to think about these concepts, but we stress that the derivative is not a fraction. Higher order derivatives are written as:

$$f^{(n)} = \frac{d^n f}{dx}$$

You will show on your challenge homework this week that if f is differentiable at a it is also continuous at a. There are however, examples of continuous functions which do not admit a derivative everywhere:

Example 2.1. Let $f(x) = |x|$, then for all $x \leq 0$ we have that $f(x) = -x$ and for all $x \geq 0$ we have that $f(x) = x$. Let $a < 0$ then:

$$f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$$

As h gets closer and closer to zero from the left and the right of a, we have that $a+h$ is still negative, hence:

$$f'(a) = \lim_{h\to 0} \frac{-(a+h)-(-a)}{h} = \lim_{h\to 0} -\frac{h}{h} = -1$$

For $a > 0$, the same argument shows that:

$$f'(a) = 1$$

We however have a problem at $a = 0$, when h approaches 0 from the right it is always positive, and when h approaches 0 from the left it is always negative. It follows that when taking the limit we have to be careful about which side we are approaching from. Proceeding with the calculation, we have that:

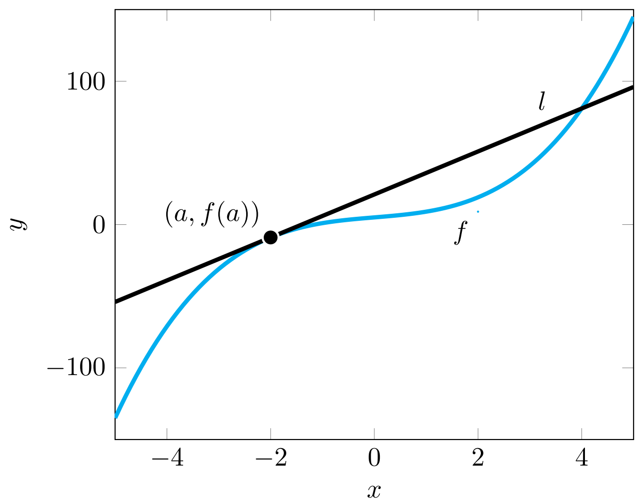
$$\lim_{h\to 0^-} \frac{f(h)-f(0)}{h} = \lim_{h\to 0^-} \frac{-h}{h} = -1$$

while:

$$\lim_{h\to 0^+} \frac{f(h)-f(0)}{h} = \lim_{h\to 0^+} \frac{h}{h} = 1$$

so $f'(0)$ does note exist, and $|x|$ is not differentiable at $x = 0$.

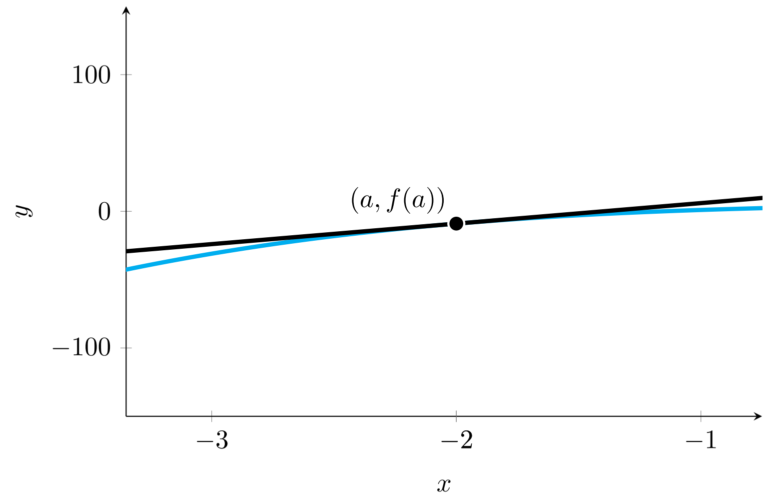
Before calculating more examples of functions, we wish to provide a geometric interpretation of the derivative. Our original motivation came from physics, where we thought of the derivative of a distance function as giving us a velocity function; but in full generality what is the derivative actually telling us? Well, the average rate of range of f on an interval $[a, b]$ is the slope of the line passing through the points $(a, f(a))$ and $(b, f(b))$. We can thus interpret the derivative of f at a as the slope of a line passing through $(a, f(a))$, but there are infinitely many such lines, parameterized by their slope, so which one is it? It turns out this line is a very special line, it is the line *tangent* to the graph of f at a. By this we mean that our line doesn't intersect our graph at $(a, f(a))$, but instead just glances off it. The following image illustrates what we mean by this:



The derivative is thus the slope of the tangent line at a point, and the slope of the tangent line is akin to a form of instantaneous rate of change. In other words, the derivative measures how the function is changing at any point. Importantly, and we will explore this topic in depth later, for h very close to zero, we can approximate $f(a+h)$ by $f(a) + h \cdot f'(a)$ This because as we zoom in, the tangent line to f at a is a good approximation for f, and this line is precisely:

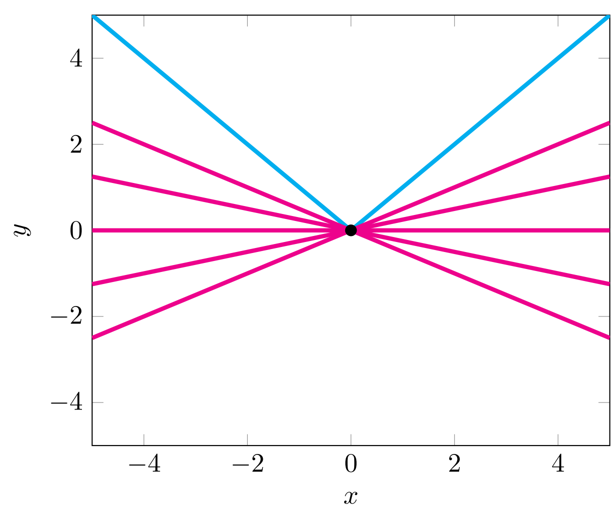
$$l(x) = f'(a) \cdot (x - a) + f(a)$$

so plugging in $x = x + h$, we obtain that that $l(a + h) = f'(a) \cdot h + f(a)$ which, as mentioned, is approximately $f(a+h)$. This is easier to see if we actually zoom in:



Note that this discussion implies that if $f'(a) > 0$ then the function f is increasing at a, if $f'(a) < 0$ then f is decreasing at a, and $f'(a) = 0$ then f is not changing at all a.

How do we then interpret Example 2.1? What does it mean for the derivative of a function, especially a continuos functions, to not exist? We have the following picture:



Essentially, there is not just one m such that $l(x) = mx$ is tangent to the graph of $|x|$, but infinitely many. The slope of the line tangent to $|x|$ at $x = 0$ is thus undefinable. This is because the how the value of $|x|$ is changing alters abruptly at $x = 0$. Indeed, for all $x < 0$, we have that $|x|$ is decreasing at a constant rate of -1, while for $x > 0$ it is increasing at a constant rate of 1. This abrupt change in the functions rate of change is demonstrated by the graph of $|x|$ being 'pointy', or not 'smooth' at $x = 0$. We can thus further interpret the derivative existing at a point as the graph of a function being smooth at that point with no sharp corners or points.

With all of the above in mind, we have a new tool, namely the derivative, and we should start calculating some examples. We first want to show the following:

Theorem 2.1. *Let f and g be functions differentiable a, and c be a real number. Then $(f+g)$ and $c \cdot f$ are differentiable functions satisfying $(f+g)'(a) = f'(a) + g'(a)$ and $(c \cdot f)'(a) = c \cdot f'(a)$. In other words, we have that as functions:*

$$(f+g)' = f' + g' \qquad \textit{and} \qquad (c \cdot f)' = c \cdot f'$$

Proof. By our limit laws for addition:

$$\begin{aligned}(f+g)'(a) &= \lim_{h\to 0}\frac{(f+g)(a+h)-(f+g)(h)}{h}\\ &= \lim_{h\to 0}\frac{f(a+h)+g(a+h)-f(h)-g(h)}{h}\\ &= \lim_{h\to 0}\frac{f(a+h)-f(h)}{h}+\frac{g(a+h)-g(h)}{h}\\ &= f'(a)+g'(a)\end{aligned}$$

as desired. By question 2 part e) on your challenge homework, we also have that:

$$\begin{aligned}(c\cdot f)'(a) &= \lim_{h\to 0}\frac{(c\cdot f)(a+h)-(c\cdot f)(h)}{h}\\ &= \lim_{h\to 0}\frac{c\cdot f(a+h)-c\cdot f(h)}{h}\\ &= \lim_{h\to 0} c\cdot\frac{f(a+h)-f(h)}{h}\\ &= c\cdot f'(a)\end{aligned}$$

as desired. □

We can now begin to tackle polynomials:

Example 2.2. Let $f(x) = x^n$ where n is a whole positive number, we want to find f'. For any h and any n, we want to know what $(x+h)^n$. This is a tricky question, but we actually only need to know what two terms of this expression look like. We have that:

$$(x+h)^n = \underbrace{(x+h)\cdots(x+h)}_{n-\text{times}}$$

Ok, so we have to multiply x with itself n times, so we know for a fact that:

$$(x+h)^n = x^n + \text{other stuff}$$

When we expand everything out, every other term will have a factor of h in it. If a term has one h in it, then it has to be of the form hx^{n-1}, and if we were to expand everything out we have n of them. Every other term has a factor of h^2 or higher in it, hence:

$$(x+h)^n = x^n + nhx^{n-1} + \text{other stuff divisible by } h^2$$

We can thus write the following:

$$\begin{aligned}f'(x) &= \lim_{h\to 0}\frac{(x+h)^n - x^n}{h}\\ &= \lim_{h\to 0}\frac{x^n + nhx^{n-1} + \text{other stuff divisible by } h^2 - x^n}{h}\\ &= \lim_{h\to 0}\frac{nhx^{n-1} + \text{other stuff divisible by } h^2 - x^n}{h}\\ &= \lim_{h\to 0} nx^{n-1} + \text{other stuff divisible by } h\\ &= nx^{n-1} + 0\end{aligned}$$

It follows that $f'(x) = nx^{n-1}$, as desired.

Example 2.3. With Theorem 2.1 and Example 2.2 we can take the derivative of any polynomial. Indeed, let:

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

then:

$$p'(x) = n \cdot a_n x^{n-1} + (n-1) \cdot a_{n-1} x^{n-2} + \cdots + a_1 + 0$$

In Example 2.2 we showed that the derivative of x^n is nx^{n-1} when n is any whole positive number. It is a fact known as the *power rule* that this is actually true for any number not equal to zero. We will not prove this in this course as it requires certain machinery we are not equipped to develop, but we do etch this rule in stone with the following theorem:

Theorem 2.2. *Let $f(x) = x^a$ for a any real non zero number. Then $f'(x) = a \cdot x^{a-1}$.*

We end the section with the following example:

Example 2.4. Let $f(x) = e^x$, then we want to show that $f'(x) = e^x$ as well. This is a fact that is true only about e, and does not hold for any other number. We will generalize this result to functions of the form a^x where a is any real positive number in the section on the chain rule. Let us begin:

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{e^{x+h} - e^x}{h} \\ &= \lim_{h \to 0} \frac{e^x e^h - e^x}{h} \\ &= \lim_{h \to 0} e^x \frac{e^h - 1}{h} \\ &= e^x \cdot \lim_{h \to 0} \frac{e^h - 1}{h} \end{aligned}$$

Where have we seen this limit before? That's right, in Example 1.14 we showed that this limit was equal to one! It follows that:

$$f'(x) = e^x$$

We give the following table of derivatives. Everything on here is fair game to ask the derivative of on an exam, so please commit this table to memory. You will spend a significant amount of your challenge homework checking this table.

Function $f(x)$	**Derivative** $f'(x)$	**Domain Notes**
c	0	All real numbers
x^n	nx^{n-1}	$x \neq 0$ if $n < 1$
e^x	e^x	All real numbers
a^x	$a^x \ln a$	$a > 0$, $a \neq 1$
$\ln x$	$\frac{1}{x}$	$x > 0$
$\log_a x$	$\frac{1}{x \ln a}$	$x > 0$, $a > 0$, $a \neq 1$
$\sin x$	$\cos x$	All real numbers
$\cos x$	$-\sin x$	All real numbers
$\tan x$	$\sec^2 x$	$x \neq \frac{\pi}{2} + n\pi$
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$	$-1 < x < 1$
$\arccos x$	$\frac{-1}{\sqrt{1-x^2}}$	$-1 < x < 1$
$\arctan x$	$\frac{1}{1+x^2}$	All real numbers
$\sinh x$	$\cosh x$	All real numbers
$\cosh x$	$\sinh x$	All real numbers

Table 1: Derivatives of common functions

2.2 The Product Rule and Quotient Rule

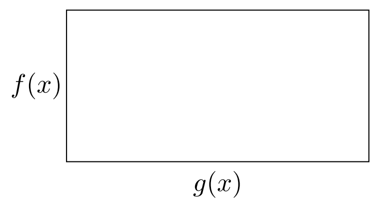
In the previous section, we were able to easily find that the derivative of a sum functions is the sum of their derivatives, and that the derivative of a real number times a function was that real number multiplied with that function. Symbolically this is stated as:

$$\frac{d}{dx}(f+g) = \frac{df}{dx} + \frac{dg}{dx} \qquad \text{and} \qquad \frac{d}{dx}(c \cdot f) = c \cdot \frac{df}{dx}$$

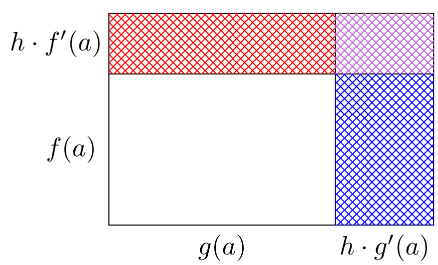
In this section we explore what happens when we try to take derivatives of products and quotients. One may initially expect that:

$$(f \cdot g)' = f' \cdot g'$$

However, if we reason this out for a moment, we will see that this doesn't quite make sense. Indeed, we can think of $f \cdot g$ as being the function such that for each a, $(f \cdot g)(a)$ is the area of the rectangle with side lengths $f(a)$ and $g(a)$:



Now if h is a number very close to zero, then we can approximate $f(a+h)$ by $f(a) + h \cdot f'(a)$, and similarly for g. It follows that the rectangle with side lengths $f(a+h)$ and $g(a+h)$ is given by:



Note that the area of the red rectangle is $h \cdot f'(a) \cdot g(a)$, the area of the blue rectangle is $h \cdot g'(a) \cdot f(a)$, and the area of the purple rectangle is $h^2 \cdot f'(a) \cdot g'(a)$¿ Ok, so now what is $\Delta(f \cdot g)$ over the interval $[a, a+h]$? Well, we have that when h is small enough:

$$\begin{aligned}(f \cdot g)(a+h) \approx &(f(a) + h \cdot f'(a)) \cdot (g(a) + h \cdot g'(a)) \\ = &f(a) \cdot g(a) + h \cdot f'(a) \cdot g(a) + h \cdot f(a) \cdot g'(a) + h^2 \cdot f'(a) \cdot g'(a)\end{aligned}$$

It follows that:

$$\Delta(f \cdot g) = (f \cdot g)(a+h) - (f \cdot g)(h) \approx h \cdot f'(a) \cdot g(a) + h \cdot f(a) \cdot g'(a) + h^2 \cdot f'(a) \cdot g'(a)$$

and so when we take the limit as $h \to 0$:

$$\begin{aligned}\lim_{h \to 0} \frac{h \cdot f'(a) \cdot g(a) + h \cdot f(a) \cdot g'(a) + h^2 \cdot f'(a) \cdot g'(a)}{h} = &\lim_{h \to 0} f'(a) \cdot g(a) + f(a) \cdot g'(a) + hf'(a)g'(a) \\ = &f'(a) \cdot g(a) + f(a) \cdot g'(a)\end{aligned}$$

This informal argument suggests to us that as we shrink h to zero, the $h^2 \cdot f'(a)g'(a)$ part of the rectangle goes to zero, and the only parts that matter are the $h \cdot f'(a) \cdot g(a)$ and $h \cdot g'(a) \cdot f(a)$ areas. This is known as the *product rule*, and we provide a correct, formal proof of it below:

Theorem 2.3. *Suppose that f and g are functions differentiable at a. Then:*

$$(f \cdot g)'(a) = f'(a) \cdot g(a) + f(a) \cdot g'(a)$$

In particular, if f and g are differentiable functions, then:

$$(f \cdot g)' = f' \cdot g + f \cdot g'$$

Proof. We have that for all $h \neq 0$, over the interval $[a, a+h]$:[11]

$$\begin{aligned}
\frac{\Delta(fg)}{\Delta x} =& \frac{f(a+h)g(a+h) - f(a)g(a)}{h} \\
=& \frac{f(a+h) \cdot g(a+h) + f(a)g(a+h) - f(a)g(a+h) + f(a+h)g(a) - f(a)g(a)}{h} \\
=& \frac{f(a+h)g(a+h) - f(a)g(a+h)}{h} + \frac{f(a)g(a_h) - f(a)g(a)}{h} \\
=& \frac{f(a+h) - f(a)}{h} \cdot g(a+h) + f(a) \cdot \frac{g(a+h) - g(a)}{h}
\end{aligned}$$

Since f and g are differentiable at a, we know that the limits:

$$\lim_{h \to 0} \frac{g(a+h) - g(a)}{h} \qquad \text{and} \qquad \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

exist. Moreover, $\lim_{h \to 0} g(a+h) = g(a)$. It follows by our limit laws, and the definition of the derivative that:

$$\begin{aligned}
\lim_{h \to 0} \frac{f(a+h)g(a+h) - f(a)g(a)}{h} =& \left(\lim_{h \to 0} \frac{f(a+h) - f(a)}{h}\right) \cdot g(a+h) + f(a) \cdot \left(\lim_{h \to 0} \frac{g(a+h) - g(a)}{h}\right) \\
=& f'(a) \cdot g(a) + f(a) \cdot g'(a)
\end{aligned}$$

as desired. □

Here is an example of the product rule in action:

Example 2.5. Let $f(x) = \sin x$ and $g(x) = \cos x$, then we want to find the derivative of $(f \cdot g)(x) = \sin x \cdot \cos x$. In other, words we want to find the derivative of:

$$h(x) = \sin x \cdot \cos x$$

With the product rule, and Table 1 we have that:

$$\begin{aligned}
h'(x) =& \cos x \cdot \cos x - \sin x \cdot \sin x \\
=& \cos^2 x - \sin^2 x
\end{aligned}$$

Example 2.6. Let $f(x) = e^x \cos x$, then:

$$f'(x) = e^x \cdot \cos x - e^x \cdot \sin x = e^x(\cos x - \sin x)$$

If we instead want to take quotients of functions, we need to employ the *quotient rule*:

Theorem 2.4. *Let f and g be differentiable at a with $g(a) \neq 0$; then the derivative of $h = f/g$ at a is given by:*

$$h'(a) = \frac{g(a) \cdot f'(a) - f(a) \cdot g'(a)}{[g(a)]^2}$$

In particular, we have that as functions:

$$h' = \frac{g \cdot f' - f \cdot g'}{g^2}$$

[11] Here $\Delta x = h$, and $\Delta(f \cdot g) = (fg)(a+h) - (f \cdot g)(a)$

One can prove this using the limit definition of the derivative, but there is a slicker using the chain rule,[12] and the power rule. You will prove the quotient rule on your challenge homework via this method.

We note that if we think of the derivative of as df/dx, the following mnemonic allows us to easily remember the chain rule: 'low d-high minus high d-low all over low squared'.

Example 2.7. Let $f(x) = e^x$ and $g(x) = x$, then:

$$\begin{aligned}\left(\frac{f}{g}\right)'(x) =& \frac{x \cdot e^x - e^x \cdot 1}{x^2} \\ =& \frac{e^x(x-1)}{x^2}\end{aligned}$$

Example 2.8. Let:

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

We recognize this immediately as a quotient of two functions, namely $h(x) = x$ and $g(x) = \ln x$. Using Table 1, and the quotient rule we have that:

$$f'(x) = \frac{\ln x - x \cdot \frac{1}{x}}{(\ln x)^2} = \frac{\ln x - 1}{(\ln x)^2}$$

2.3 The Chain Rule and Implicit Differentiation

Given two functions f and g, we have 5 ways of making new functions. We can add them $f + g$, subtract them $f - g$, multiply them $f \cdot g$, divide them f/g, and compose them $f \circ g$.[13] We know how to take derivatives of every single function operation except composition. In this section we explore how to take such derivatives. We will not be able to provide a proof, but we will attempt to justify the rule for taking derivatives of composite functions; this rule is known as the chain rule.

We suppose that f and g are functions, with g differentiable at a, and f differentiable at $g(a)$. We have that for h very close to zero:[14]

$$f(g(a+h)) \approx f(g(a) + h \cdot g'(a)) \approx f(g(a)) + h \cdot g'(a) \cdot f'(g(a))$$

It follows that:

$$f(g(a+h)) - f(g(a)) \approx h \cdot g'(a) \cdot f'(g(a))$$

If we believe all this, then the following limit is obvious

$$\lim_{h \to 0} \frac{f(g(a+h)) - f(g(a))}{h} = \lim_{h \to 0} \frac{h \cdot f'(g(a)) \cdot g'(a)}{h} = f'(g(a)) \cdot g'(a)$$

We enshrine this rule with a theorem:

Theorem 2.5. *Let f and g be functions, with g differentiable at a, and f differentiable at $g(a)$, then:*

$$(f \circ g)'(a) = f'(g(a)) \cdot g'(a)$$

On the level of functions:

$$(f \circ g)' = (f' \circ g) \cdot g'$$

[12] See next section

[13] Note that $f \circ g$ is common notation for the function defined by $f(g(x))$.

[14] The following approximations is where we are sweeping all of the hard work under the rug.

Example 2.9. Consider $f(x) = a^x$, then we can write $f(x)$ as $e^{\ln a \cdot x}$, as $e^{\ln a} = a$. If $g(x) = e^x$ and $h(x) = \ln a \cdot x$, then $f(x) = g(h(x))$, and so by the chain rule:

$$f'(x) = g'(h(x)) \cdot h'(x) = e^{\ln a \cdot x} \cdot \ln a = \ln a \cdot a^x$$

which is our general exponential derivative rule from Table 1.

Example 2.10. Let $h(x) = e^{\sin x}$, then we have that $h(x) = f(g(x))$, where $f(x) = e^x$, and $g(x) = \sin x$. Taking a derivative of $h(x)$ we have that by the chain rule:

$$h'(x) = f'(g(x)) \cdot g'(x)$$

We know that $g'(x) = \cos x$, and that $f'(x) = e^x$, then $f'(g(x)) = e^{\sin x}$. It follows that:

$$h'(x) = e^{\sin x} \cdot \cos x$$

This specific rule is, in my opinion, best remembered using the fractional notation of the derivative. Indeed, we have that $g'(a) = dg/dx|_a$, while $f'(g(a)) = df/dx|_{g(a)}$. If we think of the composition $f \circ g$ as '$f \circ g$ being a function of g', then we can write $df/dx|_{g(a)}$ as $d(f \circ g)/dg|_a$, where by $d(f \circ g)/dg$ we mean $f' \circ g$.[15] With $h = f \circ g$:

$$\left.\frac{dh}{dx}\right|_a = \left.\frac{df}{dg}\right|_a \cdot \left.\frac{dg}{dx}\right|_a$$

On the level of functions:

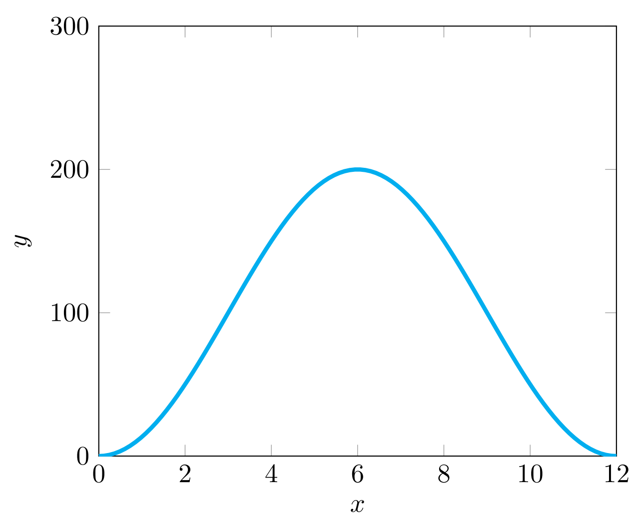
$$\frac{dh}{dx} = \frac{d(f \circ g)}{dg} \cdot \frac{dg}{dx}$$

Abstractly thinking of one function as a function of another function is hard, so let us come up with a reasonable example.

Example 2.11. Suppose the amount of vegetation in a meadow measured in kilograms is given by function of t months:

$$v(t) = -100 \cos\left(\frac{\pi}{6} t\right) + 100$$

where we interpret $t = 0$ as being the start of January. Note this attempts to depict a realistic picture for how much vegetation a meadow would have in any given month, peaking in the spring and summer months:



[15]This is the function defined by $(f' \circ g)(x) = f'(g(x))$.

Now further suppose that the population of rabbits in the meadow is a function of the available vegetation:

$$p(v) = 10 + \left(1000^{1/200}\right)^{v} \tag{2.4}$$

Note this means that when $v = 200$, the maximum amount of vegetation in the meadow, we will have a population of 1010 rabbits. We can calculate how the population of rabbits changes with vegetation, and how the vegetation changes with time; the chain rule says that this is enough to know how the population changes with time.

We see that:

$$\frac{dp}{dv} = \frac{1}{200}\ln(1000) \cdot \left(1000^{1/200}\right)^{v}$$

while:

$$\frac{dv}{dt} = \frac{100 \cdot \pi}{6} \cdot \sin\left(\frac{\pi}{6}t\right)$$

Now since p is a function of v and v is a function of t, we can take a derivative of p with respect to t, which measures how p changes with respect to t. The chain rule states that this derivative is given by:

$$\begin{aligned}\frac{dp}{dt} =& \frac{dp}{dv} \cdot \frac{dv}{dt} \\ =& \left(\frac{1}{200}\ln(1000) \cdot \left(1000^{1/200}\right)^{v}\right) \cdot \left(\frac{100 \cdot \pi}{6} \cdot \sin\left(\frac{\pi}{6}t\right)\right) \\ =& \frac{\pi}{12}\left(1000^{1/200}\right)^{v} \cdot \sin\left(\frac{\pi}{6}t\right)\end{aligned}$$

We can replace v with the (2.4) to obtain:

$$\frac{dp}{dt} = \frac{\pi}{12}\left(1000^{1/200}\right)^{-100\cos\left(\frac{\pi}{6}t\right)+100} \cdot \sin\left(\frac{\pi}{6}t\right)$$

Example 2.12. A rocket ships distance from earth is given in kilometers by:

$$r(t) = \ln(t+1)$$

The gravitational force that the Earth exerts on the rocket ship is given as a function of r:

$$F(r) = \frac{k}{r^2}$$

where k a constant relating the strength of gravity and the masses of the earth and the rocket ship. Using the chain rule, we can find out the force the Earth exerts on the rocket ship changes with time. Indeed, we have that:

$$\frac{dF}{dr} = -\frac{2k}{r^3}$$

while:

$$\frac{dr}{dt} = \frac{1}{t+1}$$

It follows that:

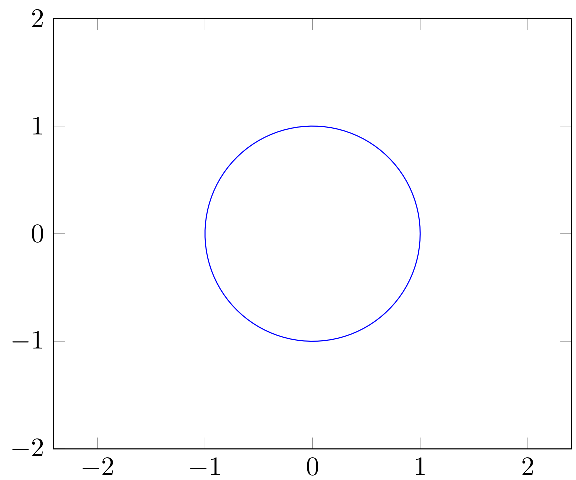
$$\begin{aligned}\frac{dF}{dt} =& \frac{dF}{dr} \cdot \frac{dr}{dt} \\ =& -\frac{2k}{r^3} \cdot \frac{1}{t+1} \\ =& -\frac{2k}{(\ln(t+1))^3} \cdot \frac{1}{t+1}\end{aligned}$$

A particularly apt application of the chain rule is something called implicit differentiation. We explain this as follows: if we have a function $f(x)$, then when we graph a function we set the y coordinate equal to $f(x)$. In particular, this motivates using df/dx and dy/dx interchangeably to refer to the derivative, . However, sometimes we care about graphs that aren't quite functions, but instead a relation between x and y, i.e. instead of $y = f(x)$, we have something like $xy = 1$, or $x^2 + y^2 = 1$. We want to be able to calculate the tangent line to curves in the plane of this form. We demonstrate this via example:

Example 2.13. Suppose that we have the following curve:

$$x^2 + y^2 = 1 \tag{2.5}$$

The set of points in the plain (x, y) which satisfy this relation forms a circle:



Suppose we want to find the slope of the tangent line at $(\sqrt{2}/2, \sqrt{2/2})$. Note that this is point lies on the circle as:

$$\left(\frac{\sqrt{2}}{2}\right)^2 + \left(\frac{\sqrt{2}}{2}\right)^2 = \frac{2+2}{4} = 1$$

Well we simply take the derivative of both sides of (2.5)! We know that:

$$\frac{d}{dx}(1) = 0 \qquad \text{and} \qquad \frac{d}{dx}x^2 = 2x$$

but what about y^2? Well, we treat y as if *were a function of* x, and apply the chain rule to get:

$$\frac{d}{dx}(y^2) = 2y \cdot \frac{dy}{dx}$$

Putting it all together we obtain that:

$$2x + 2y \cdot \frac{dy}{dx} = 0$$

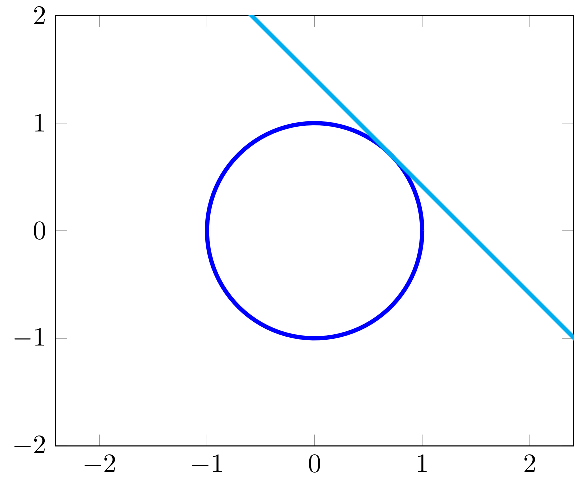
Now we can plug in $x = \sqrt{2}/2$ and $y = \sqrt{2}/2$ to solve for dy/dx:

$$\sqrt{2} + \sqrt{2} \cdot \frac{dy}{dx} = 0 \Rightarrow \sqrt{2} \cdot \frac{dy}{dx} = -\sqrt{2} \Rightarrow \frac{dy}{dx} = -1$$

It follows that our tangent line $l(x)$ is given by:

$$l(x) = -1(x - \sqrt{2}/2) + \sqrt{2}/2$$

We graph this to make sure:



In particular, we can write dy/dx as the following function of both x and y:

$$\frac{dy}{dx} = -\frac{x}{y}$$

Example 2.14. Now suppose we want to differentiate the function:

$$y^3 - x \sin y = 8$$

Then we have that by the chain rule:

$$\frac{d}{dx} y^3 = 3y^2 \cdot \frac{dy}{dx}$$

and by the product rule:

$$\frac{d}{dx}(x \sin y) = \sin y + x \cos y \frac{dy}{dx}$$

hence:

$$3y^2 \cdot \frac{dy}{dx} - \sin y + x \cos y \frac{dy}{dx} = 0$$

We move all the terms that contain dy/dx to one side to get:

$$3y^2 \cdot \frac{dy}{dx} + x \cos y \frac{dy}{dx} = -\sin y$$

We can then factor out dy/dx from each term to obtain that:

$$\frac{dy}{dx} = \frac{-\sin y}{3y^2 + x \cos y}$$

Note it is not always possible to write this solely as a function of x.

2.4 Inverse Function Theorem

let f be a function, then we say that f has an inverse, denoted f^{-1} if:

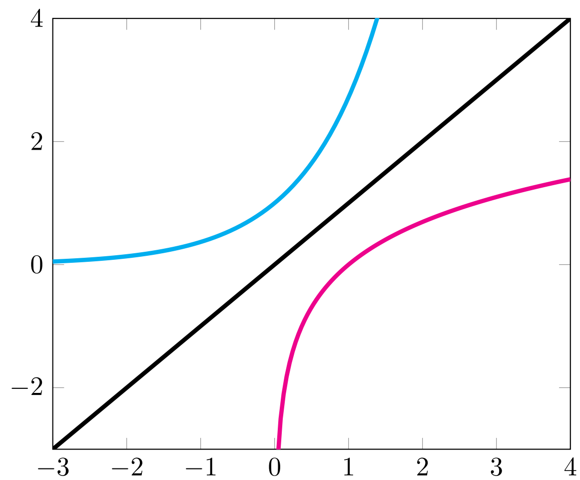
$$f(f^{-1}(x)) = f^{-1}(f(x)) = x$$

Importantly, if such an inverse exists, then the graph of f^{-1} is the graph of f flipped over the line $y = x$ in the plane. The domain of the inverse function, is always the range of the original function, but the range is not always the domain. We will see an example of this shortly.

Example 2.15. If $f(x) = 2^x$, then it's inverse is $\log_2(x)$ essentially by definition. This actually true of any function of the form a^x, i.e. $\log_a x$ is it's inverse. We can see this because for all x:

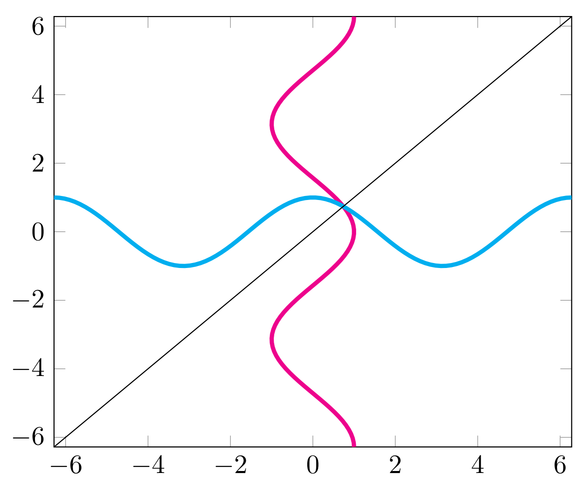
$$2^{\log_2 x} = x \qquad \text{and} \qquad \log_2 2^x = x$$

We graph these functions to demonstrate the flipping:

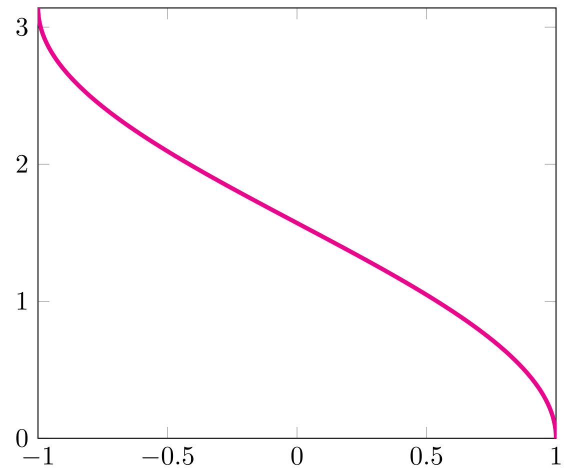


The blue plot is 2^x, and the magenta curve is $\log_2 x$. The black line is $y = x$.

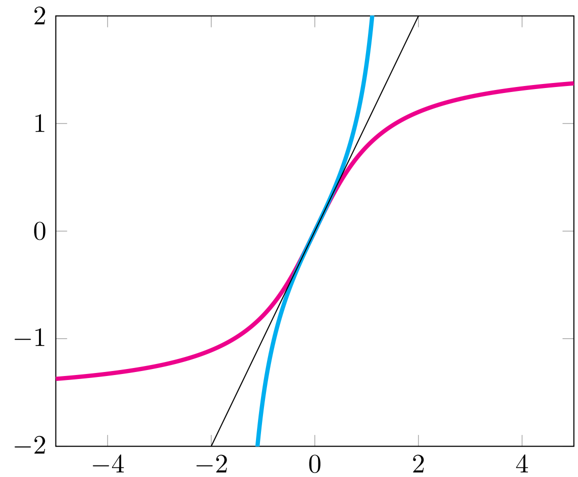
Example 2.16. The trigonometric functions sin, cos and tan have inverse functions given by arcsin, arccos and arctan. These are sometimes also referred to by $\sin^{-1}$, $\cos^{-1}$, and $\tan^{-1}$. Note that domain of $\sin^{-1}$ and $\cos^{-1}$ is $[-1, 1]$, but the range is only $[-\pi/2, \pi/2]$ for $\sin^{-1}$ and $[0, \pi]$ for $\cos^{-1}$. This is because if we reflect $\cos x$ (or $\sin x$) over the line $y = x$ we get the following picture:



The magenta is our supposed inverse, but this clearly would not function the vertical line test, i.e. it would have many outputs. If instead we restrict our range to $[0, \pi]$, we get:



which does pass the vertical line test. I suggest you map out $\sin x$ and $\sin^{-1} x$ in a similar fashion on Desmos. We graph the situation for $\tan x$ and $\tan^{-1} x$ here as well, as $\tan^{-1} x$ is defined everywhere:



which looks as expected.

Example 2.17. Another important example is every linear function. Indeed, suppose that:

$$f(x) = mx + b$$

then we claim that:

$$f^{-1}(x) = \frac{1}{m}(x - b)$$

Indeed:

$$\begin{aligned} f(f^{-1}(x)) =& f\left(\frac{1}{m}(x - b)\right) \\ =& m \cdot \frac{1}{m}(x - b) + b \\ =& x - b + b \\ =& x \end{aligned}$$

similarly:

$$\begin{aligned}f^{-1}(f(x)) =& f^{-1}(mx+b)\\ =&\frac{1}{m}(mx+b-b)\\ =&\frac{1}{m}mx\\ =&x\end{aligned}$$

Now how can we take derivatives of inverse functions? Given that the graph of an inverse function is the graph of the original function just flipped over the line $y = x$, it appears that the inverse should be differentiable since it looks like we can draw it smoothly, but what exactly is the derivative?

Theorem 2.6. *Suppose that f is a differentiable function, and $f'(a) \neq 0$. Then if there exists an inverse function f^{-1}, the derivative at $x = f(a)$ is given by:*

$$f^{-1'}(f(a)) = \frac{1}{f'(a)}$$

Proof. Since the graph of f^{-1} is smooth, we assume that it is differentiable. Now we have that:

$$f^{-1}(f(x)) = x$$

The derivative of x is one, hence:

$$\frac{d}{dx}\left(f^{-1}(f(x))\right) = 1$$

The chain rule tells us that the left hand side is given by:

$$\frac{d}{dx}\left(f^{-1}(f(x))\right) = f^{-1'}(f(x)) \cdot f'(x)$$

At $x = a$ this becomes:

$$f^{-1'}(f(a)) \cdot f'(a) = 1 \Rightarrow f^{-1'}(f(a)) = \frac{1}{f'(a)}$$

as desired. □

We immediately have the following example:

Example 2.18. We want to know what the derivative of $\ln x$ at $x = a$ is. Well, if $a > 0$, which is the only place that $\ln x$ is defined, we can write a as $e^{\ln a}$. Now set $b = \ln a$, then we have that by the inverse function theorem:

$$\left(\frac{d}{dx}\ln\right)(e^b) = \frac{1}{e^b}$$

because the derivative of e^x is e^x. Putting in $\ln a$ for b, we get that:

$$\left(\frac{d}{dx}\ln\right)(a) = \frac{1}{a}$$

as $e^{\ln a} = a$. It follows that:

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

Now what about $\log_c x$ for some real number $c > 0$? Well:

$$\log_c x = \frac{\ln x}{\ln c}$$

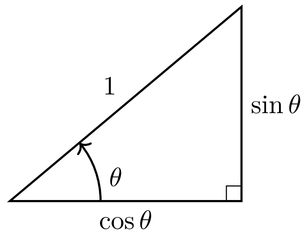
so taking a derivative gives us:

$$\frac{d}{dx}\log_c x = \frac{1}{x \cdot \ln c}$$

Example 2.19. Let us derive the formula for the derivative of $\arcsin x$ as well. If $-1 < a < 1$, we have that we can write a as $\sin\theta$ for some θ in the interval $(-\pi/2, \pi/2)$. It follows that:

$$\left(\frac{d}{dx}\arcsin\right)(\sin\theta) = \frac{1}{\cos(\theta)}$$

We need to figure out what $\cos(\theta)$ is. Consider the following diagram:



Now since $\sin\theta = a$, and:

$$\sin^2\theta + \cos^2\theta = 1 \tag{2.6}$$

we must have that:

$$\cos\theta = \sqrt{1 - \sin^2\theta} = \sqrt{1 - a^2}$$

hence:

$$\left(\frac{d}{dx}\arcsin\right)(a) = \frac{1}{\sqrt{1 - a^2}}$$

Since this holds for all a we have that:

$$\frac{d}{dx}\arcsin x = \frac{1}{\sqrt{1 - x2}}$$

Example 2.20. We want to derive the formula of the derivative of $\arccos x$. If $-1 < a < 1$ we have that we can write a as $\cos\theta$ for some θ in the interval $(-\pi, \pi)$. The inverse function theorem gives us:

$$\left(\frac{d}{dx}\arccos\right)(\cos\theta) = \frac{1}{-\sin\theta}$$

Now what is $-\sin\theta$? By the same argument as in the previous example, we have that:

$$\sin^2\theta + \cos^2\theta = 1$$

hence:

$$\sin\theta = \sqrt{1 - \cos^2\theta} = \sqrt{1 - a^2}$$

it follows that:

$$-\sin\theta = -\sqrt{1 - a^2}$$

Therefore:

$$\left(\frac{d}{dx}\arccos\right)(a) = \frac{-1}{\sqrt{1 - a^2}}$$

Since this holds for all $-1 < a < 1$:

$$\frac{d}{dx}\arccos x = \frac{-1}{\sqrt{1 - a^2}}$$

Example 2.21. Finally, we want to derive a formula for $\arctan x$. For any x we can write it as $\tan\theta$ for some θ in the interval $(-\pi/2, \pi/2)$. We can thus write:

$$\left(\frac{d}{dx}\arctan\right)(\tan\theta) = \frac{1}{\sec^2\theta}$$

Now, dividing equation (2.6) by $\cos^2\theta$ gives us:

$$\tan^2\theta + 1 = \sec^2\theta$$

hence we have that $\sec^2\theta = 1 + a^2$ as $a = \tan\theta$. It follows that:

$$\left(\frac{d}{dx}\arctan\right)(a) = \frac{1}{1+a^2}$$

Since this holds for all real numbers a, we have that:

$$\frac{d}{dx}\arctan x = \frac{1}{1+a^2}$$

Week III: Examples

3.1 Examples: Product Rule, Quotient Rule, Chain Rule

Example 3.1. In this example we go over how to calculate the derivative of:

$$f(x) = \frac{e^x}{\tan x}$$

Since f is a *quotient* of two functions, we should use the quotient rule. We know that:

$$\frac{d}{dx}e^x = e^x \qquad \text{and} \qquad \frac{d}{dx}\tan x = \sec^2 x$$

hence the quotient rule tells us that:

$$\begin{aligned}\frac{df}{dx} &= \frac{\tan x \frac{d}{dx}e^x - e^x \frac{d}{dx}\tan x}{\tan^2 x} \\ &= \frac{e^x \tan x - e^x \sec^2 x}{\tan^2 x}\end{aligned}$$

Example 3.2. What if we want to find the tangent line to:

$$f(x) = \frac{e^x}{\tan x}$$

at $x = \pi/4$? Well the derivative tells us the slope of the tangent line at any point, and we know that the tangent line has to pass through the point $(\pi/4, f(\pi/4))$ so point slope form tells us that the tangent line should be given by:

$$y - f(\pi/4) = \frac{df}{dx}(\pi/4)(x - \pi/4)$$

We just need to find $f(\pi/4)$ and $(df/dx)(\pi/4)$. Note that $\tan\pi/4 = \sin(\pi/4)/\cos(\pi/4) = 1$. It follows that:

$$f(\pi/4) = e^{\pi/4}$$

which we can't simplify any further. Now:

$$\sec^2(\pi/4) = \frac{1}{\cos^2(\pi/4)} = \frac{1}{(\sqrt{2}/2)^2} = \frac{1}{1/2} = 2$$

so:

$$\frac{df}{dx}(\pi/4) = \frac{e^{\pi/4} - 2e^{\pi/4}}{1} = -e^{\pi/4}$$

so our line is given by:

$$y - e^{\pi/4} = -e^{\pi/4}(x - \pi/4)$$

Example 3.3. What if we want to find the tangent line to:

$$g(x) = \cos x \cdot \sin x$$

at $x = \pi/2$? Well we need to find dg/dx first; in this case we should use the product rule because $g(x)$ is a *product* of two functions. The product rule tells us that:

$$\begin{aligned}\frac{dg}{dx} =& \frac{d}{dx}(\cos x)\sin x + \cos x \cdot \frac{d}{dx}(\sin x)\\ =& -\sin^2 x + \cos^2 x\end{aligned}$$

Plugging in $\pi/2$ we have that:

$$\begin{aligned}\frac{dg}{dx}(\pi/2) =& -\sin^2(\pi/2) + \cos^2(\pi/2)\\ =& -1 \cdot 1 + 0\\ =& -1\end{aligned}$$

Meanwhile $g(\pi/2) = 0$ because $\cos(\pi/2) = 0$. It follows that our line is given by:

$$y = -(x - \pi/2)$$

Example 3.4. What if we are given two functions:

$$f(x) = \sin x \qquad \text{and} \qquad g(x) = \ln x$$

then we can get new functions:

$$h(x) = f(g(x)) = f(\ln x) = \sin(\ln x) \qquad \text{and} \qquad l(x) = g(f(x)) = g(\sin x) = \ln(\sin x)$$

Lets find tangent lines to the functions $h(x)$ and $l(x)$ at $x = e^{\pi/4}$, and $x = \pi/4$ respectively. The the chain rule tells us that:

$$\frac{dh}{dx} = \frac{df(g)}{dx} = \frac{df(g)}{dg} \cdot \frac{dg}{dx}$$

Replacing this functions, we obtain that:

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

Here the notation $d\sin(\ln x)/d\ln x$, just means take the derivative of $\sin \ln x$, while pretending that $\ln x$ is the variable. It might help to think of $\ln x$ as being equal to y, then $d\sin(\ln x)/d\ln x$ is just $d\sin(y)/dy$ which is $\cos y$. Plugging $\ln x$ back in for for y tells us that:

$$\frac{d\sin(\ln x)}{d\ln x} = \cos(\ln x)$$

It follows that:

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

Plugging in $e^{\pi/4}$ we get:

$$\begin{aligned}\frac{dh}{dx}(e^{\pi/4}) &= \cos\Big(\ln\Big(e^{\pi/4}\Big)\Big)\cdot\frac{1}{e^{\pi/4}}\\ &= \cos(\pi/4)\cdot\frac{1}{e^{\pi/4}}\\ &= \frac{\sqrt{2}}{2e^{\pi/4}}\end{aligned}$$

Now since:

$$h(e^{\pi/4}) = \sin(\pi/4) = \frac{\sqrt{2}}{2}$$

it follows that the tangent line is given by:

$$y - \frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{2e^{\pi/4}}(x - e^{\pi/4})$$

For $l(x)$, we take the derivative the same way:

$$\begin{aligned}\frac{dl}{dx} = \frac{d\ln(\sin x)}{dx} &= \frac{d\ln(\sin x)}{d\sin x}\cdot\frac{d\sin x}{dx}\\ &= \frac{1}{\sin x}\cdot\cos x\\ &= \cot x\end{aligned}$$

Plugging in $\pi/4$ we have that:

$$\frac{dl}{dx}(\pi/4) = \cot(\pi/4) = 1$$

while:

$$l(\pi/4) = \ln\Big(\sqrt{2}/2\Big)$$

so our tangent line is:

$$y - \ln\Big(\sqrt{2}/2\Big) = x - \pi/4$$

Example 3.5. What if we have a function that is a composition of three functions? I.e.

$$h(x) = \ln x \qquad g(x) = \sin x \qquad f(x) = e^x$$

then set:

$$l(x) = h(g(f(x))) = h(g(e^x)) = h(\sin(e^x)) = \ln(\sin(e^x))$$

How do we take the derivative of $l(x)$? Well the chain rule says that:

$$\frac{dl}{dx} = \frac{d\ln(\sin(e^x))}{d\sin(e^x)}\cdot\frac{d\sin(e^x)}{dx}$$

We know the derivative of $\ln(y)$ with respect to y, it is just:

$$\frac{d\ln(y)}{dy} = \frac{1}{y}$$

Therefore:

$$\frac{d\ln(\sin(e^x))}{d\sin(e^x)} = \frac{1}{\sin(e^x)}$$

But what about $d(\sin(e^x))/dx$? Well we just apply chain rule again! Indeed, by the chain rule:

$$\begin{aligned}\frac{d\sin(e^x)}{dx} &= \frac{d\sin(e^x)}{de^x} \cdot \frac{de^x}{dx} \\ &= \cos(e^x) \cdot e^x\end{aligned}$$

It follows that:

$$\begin{aligned}\frac{dl}{dx} &= \frac{d\ln(\sin(e^x))}{d\sin(e^x)} \cdot \frac{d\sin(e^x)}{de^x} \cdot \frac{de^x}{dx} \\ &= \frac{1}{\sin(e^x)} \cdot \cos(e^x) \cdot e^x\end{aligned}$$

In general we have that:

$$\begin{aligned}\frac{dl}{dx} &= \frac{dh(g(f))}{dx} \\ &= \frac{dh(g(f))}{dg(f)} \cdot \frac{dg(f)}{dx} \\ &= \frac{dh(g(f))}{dg(f)} \cdot \frac{dg(f)}{df} \cdot \frac{df}{dx}\end{aligned}$$

In other words, we just keep applying the chain rule until we get something at the end that we can make sense of, i.e. we hit product rule, or quotient rule, or a basic derivative.

3.2 Examples: More Chain Rule, Implicit Differentiation

Example 3.6. Let us consider $f(x) = \ln(\tan x)$, we want to find it's derivative. We should use the chain rule:

$$\begin{aligned}\frac{df}{dx} &= \frac{d\ln(\tan x)}{d\tan x} \cdot \frac{d\tan x}{x} \\ &= \frac{1}{\tan x} \cdot \sec^2 x \\ &= \frac{1}{\frac{\sin}{\cos}} \frac{1}{\cos^2} \\ &= \frac{1}{\sin x \cos x} \\ &= \sec x \csc x\end{aligned}$$

Or you can just leave this as:

$$\frac{df}{dx} = \frac{\sec^2}{\tan x}$$

Example 3.7. Let us find the tangent line to:

$$g(x) = e^{\cos x}$$

at $x = \pi/4$. Let's take a derivative (using the chain rule!):

$$\begin{aligned}\frac{dg}{dx} &= \frac{de^{\cos x}}{d\cos x} \cdot \frac{d\cos x}{dx} \\ &= e^{\cos x} \cdot (-\sin x) \\ &= -e^{\cos x} \sin x\end{aligned}$$

Since the derivative gives us the slope of the tangent line, we have that our slope is:

$$\begin{aligned}\frac{df}{dx}(\pi/4) &= -\,e^{\cos(\pi/4)}\sin(\pi/4)\\ &= -\,e^{\sqrt{2}/2}(\sqrt{2}/2)\\ &= -\,\frac{\sqrt{2}}{2}e^{\sqrt{2}/2}\end{aligned}$$

while:

$$g(\pi/4) = e^{\sqrt{2}/2}$$

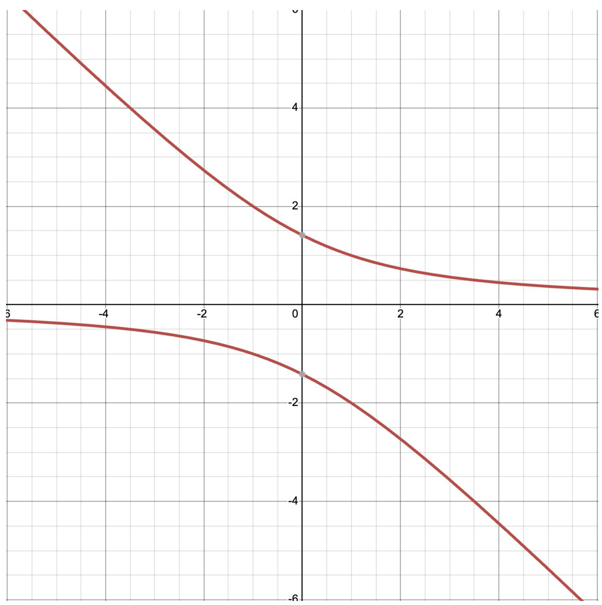
hence the tangent line is given by:

$$y - e^{\sqrt{2}/2} = -\frac{\sqrt{2}}{2}e^{\sqrt{2}/2}(x - \pi/4)$$

Example 3.8. Sometimes, we are not given a function to graph, but instead an equation:

$$xy + y^2 = 2$$

The graph of this equation is all the points (x, y) in the plane that satisfy the above relation. This is given by:



Clearly, there are tangent lines to this graph because everything looks smooth, but how do we find them? The trick to pretend that y is a function of x, (i.e. $y = f(x)$), then dy/dx will give us the slope of the tangent line at a point. Taking a derivative of both sides gives:

$$\frac{d}{dx}(xy + y^2) = \frac{d}{dx}2$$

Note that $d/dx(2)$ is zero because 2 is a constant. The derivative distributes over sums:

$$\frac{d}{dx}(xy) + \frac{d}{dx}(y^2) = 0$$

The product rule tells us that:

$$\frac{d}{dx}(xy) = \frac{dx}{dx}y + x\frac{d}{yx} = 1 + x\frac{dy}{dx}$$

We can't reduce dy/dx to anything simpler because that's what we are trying to solve for! For the other term, the chain rule tells us that:

$$\begin{aligned}\frac{dy^2}{dx} &= \frac{dy^2}{dy} \cdot \frac{dy}{x} \\ &= 2y \cdot \frac{dy}{dx}\end{aligned}$$

It follows that:

$$1 + x\frac{dy}{dx} + 2y \cdot \frac{dy}{dx} = 0$$

To solve for dy/dx, we move the 1 to the other side:

$$x\frac{dy}{dx} + 2y \cdot \frac{dy}{dx} = -1$$

Factor out dy/dx to get:

$$(x + 2y)\frac{dy}{dx} = -1$$

We divide by $(x + 2y)$ to find that:

$$\frac{dy}{dx} = \frac{-1}{x + 2y}$$

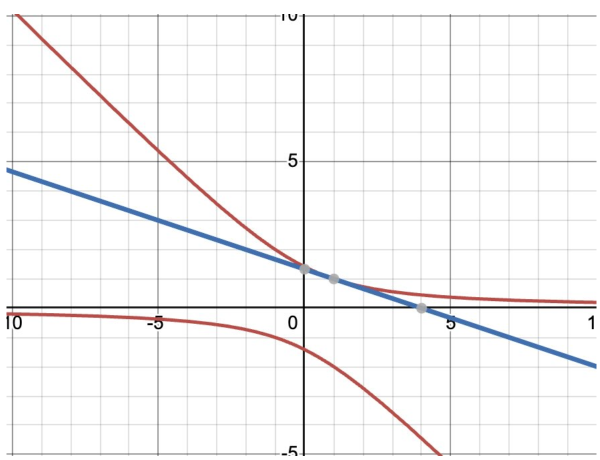
Now consider the point $(1, 1)$, this on the graph of the equation because $1 \cdot 1 + 1^2 = 2$. The slope of the tangent line is then given by plugging $x = 1$ $y = 1$ into dy/dx:

$$\frac{dy}{dx}(1, 1) = \frac{-1}{1 + 2} = -\frac{-1}{3}$$

The tangent line is then given by:

$$y - 1 = -\frac{1}{3}(x - 1)$$

Graphing this we get:



Example 3.9. We want to do the same thing as in the previous example for the following equation:

$$e^{xy} = x + y$$

i.e. we want to find dy/dx for the above equation. We begin by taking d/dx of both sides:

$$\frac{d}{dx}(e^{xy}) = \frac{d}{dx}(x+y)$$

The right hand side is easy, it just evaluates to:

$$\frac{d}{dx}(x+y) = 1 + \frac{dy}{dx}$$

The left hand side is a hair more complicated. Using the chain rule and the product rule:

$$\begin{aligned}
\frac{d}{dx}(e^{xy}) =& \frac{de^{xy}}{dx} \\
=& \frac{de^{xy}}{d(xy)} \cdot \frac{d(xy)}{dx} \\
=& e^{xy} \cdot (\frac{dx}{dx} + x \cdot \frac{dy}{dx}) \\
=& e^{xy} \cdot (1 + x\frac{dy}{dx}) \\
=& e^{xy} + xe^{xy}\frac{dy}{dx}
\end{aligned}$$

It follows that:

$$e^{xy} + xe^{xy}\frac{dy}{dx} = 1 + \frac{dy}{dx}$$

We subtract dy/dx and e^{xy} from both sides to get:

$$xe^{xy}\frac{dy}{dx} - \frac{dy}{dx} = 1 - e^{xy}$$

Pull out dy/dx to get:

$$(xe^{xy} - 1)\frac{dy}{dx} = 1 - e^{xy}$$

hence:

$$\frac{dy}{dx} = \frac{1 - e^{xy}}{xe^{xy} - 1}$$

Example 3.10. Let us now consider the equation:

$$\ln(\cos(e^y)) = xy$$

We want to find dy/dx using implicit differentiation. We take the derivative of both sides; the right hand side gives by the product rule:

$$\frac{d}{dx}(xy) = y + x\frac{dy}{dx}$$

The left hand side requires applying the chain rule multiple times:

$$\begin{aligned}
\frac{d}{dx}(\ln(\cos(e^y)) =& \frac{d\ln(\cos(e^y))}{d\cos(e^y)} \cdot \frac{d\cos(e^y)}{dx} \\
=& \frac{d\ln(\cos(e^y))}{d\cos(e^y)} \cdot \frac{d\cos(e^y)}{de^y} \cdot \frac{de^y}{dx} \\
=& \frac{d\ln(\cos(e^y))}{d\cos(e^y)} \cdot \frac{d\cos(e^y)}{de^y} \cdot \frac{de^y}{dy} \cdot \frac{dy}{dx} \\
=& \frac{1}{\cos e^y} \cdot (-\sin(e^y)) \cdot e^y \cdot \frac{dy}{dx} \\
=& -\tan(e^y) \cdot e^y \cdot \frac{dy}{dx}
\end{aligned}$$

Equating the left hand side and the right hand side:

$$-\tan(e^y) \cdot e^y \cdot \frac{dy}{dx} = y + x\frac{dy}{dx}$$

Moving all dy/dx terms to the left hand side, and multiplying throughout by -1 we obtain:

$$\tan(e^y) \cdot e^y \cdot \frac{dy}{dx} + x\frac{dy}{dx} = -y$$

Pulling out dy/dx:

$$\begin{aligned}(\tan(e^y) \cdot e^y + x)\frac{dy}{dx} &= -\, y \\ \frac{dy}{dx} &= \frac{-y}{\tan(e^y) \cdot e^y + x}\end{aligned}$$

3.3 Examples: Physical Chain Rule and Logarithmic Differentiation

Example 3.11. The gravitational pull of the Earth on a rocket ship is given by:

$$F(r) = \frac{k}{r^2}$$

where r is the distance between the Earth and the rocket ship in kilometers, and $F(r)$ is in Newtons. The distance of the rocket ship from earth in kilometers is given as a function of time:

$$r(t) = \ln(t+1)$$

where t has units seconds. We want to know how the gravitational pull of Earth on the rocket ship is changing with time, and to do that we need to find the derivative of F with respect to time t. The chain rule states that we have:

$$\frac{dF}{dt} = \frac{dF}{dr} \cdot \frac{dr}{dt}$$

Note that dF/dr has units newtons per kilometer, and dr/dt has units kilometers per second, so their product has units Newtons per second, so the chain rule makes physical sense here.

To actually calculate, we see that:

$$\frac{dF}{dr} = \frac{d}{dr}(kr^{-2})$$

so the power rule tells us that:

$$\frac{dF}{dr} - 2kr^{-3} = \frac{-2k}{r^3}$$

while:

$$\begin{aligned}\frac{dr}{dt} &= \frac{d\ln(t+1)}{d(t+1)} \cdot \frac{d(t+1)}{dt} \\ &= \frac{1}{t+1} \cdot 1 \\ &= \frac{1}{t+1}\end{aligned}$$

Multiplying the two together, we get that:

$$\frac{dF}{dt} = \frac{-2k}{r^3} \cdot \frac{1}{t+1}$$

Since $r(t) = \ln(t+1)$ we have:

$$\frac{dF}{dt} = \frac{-2k}{(\ln(t+1))^3} \cdot \frac{1}{t+1}$$

Physically, at $t = e - 1$ seconds, we have that the rocket ship is 1 km from the earth, and moving at a speed of $1/e$ kilometers per second. At 1 km, the gravity the Earth exerts on the rocket ship is changing at a rate of $-2k$ Newtons per kilometer. It follows that at $t = e - 1$, the gravity the Earth exerts on the rocket ship is changing at a rate of $-2k/e$ Newtons per second.

Example 3.12. We want to find the tangent line to:

$$f(x) = \cos\left(x^2\right)$$

at $x = \sqrt{\pi/4}$. We take a derivative and use the chain rule and power rule:

$$\begin{aligned} \frac{df}{dx} =& \frac{d\cos\left(x^2\right)}{dx^2} \cdot \frac{dx^2}{dx} \\ =& -2x\sin\left(x^2\right) \end{aligned}$$

The slope of our tangent line is thus:

$$\begin{aligned} \frac{df}{dx}(\sqrt{\pi/4}) =& -2(\sqrt{\pi/4}) \cdot \sin(\pi/4) \\ =& -\frac{2\sqrt{\pi}}{2} \cdot \frac{\sqrt{2}}{2} \\ =& -\frac{\sqrt{2\pi}}{2} \end{aligned}$$

while:

$$f(\pi/4) = \sqrt{2}/2$$

so the tangent line is given by:

$$y - \frac{\sqrt{2}}{2} = -\frac{\sqrt{2\pi}}{2}(x - \sqrt{\pi/4})$$

Example 3.13. Suppose that:

$$h(x) = 4f(x) + \frac{g(x)}{2}$$

then the derivative rules tell us that:

$$\begin{aligned} \frac{dh}{dx} =& \frac{d}{dx}\left(4f(x) + \frac{g(x)}{2}\right) \\ =& 4 \cdot \frac{df}{dx} + \frac{1}{2}\frac{dg}{dx} \end{aligned}$$

Example 3.14. In the next two examples we explore logarithmic differentiation. Consider the following function:

$$f(x) = x^x$$

How do we take it's derivative? Well we can't apply the power rule, as it is not actually a polynomial because the exponent isn't a constant, and we can't apply the exponential rule because the base isn't constant. What we can do though, is let $y = f(x)$, and use implicit differentiation. Indeed, if we have:

$$y = x^x$$

and we take the natural log of both sides, we obtain that:

$$\begin{aligned}\ln y =& \ln(x^x)\\ =& x \cdot \ln x\end{aligned}$$

Taking a derivative of both sides, we obtain:

$$\frac{1}{y} \cdot \frac{dy}{dx} = \ln(x) + 1$$

Multiplying throughout by y, we have that:

$$\frac{dy}{dx} = y \cdot (\ln(x) + 1)$$

but $y = x^x$, hence:

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

Example 3.15. We can also use logarithmic differentiation to avoid using the quotient rule. Indeed, consider:

$$f(x) = \frac{(x^2+1)^5 \cdot \sqrt{\sin x}}{(x^7-4)^2}$$

If we let $y = f(x)$ and then take the natural log both sides again, we get using the log rules:

$$\begin{aligned}\ln y =& \ln\left(\frac{(x^2+1)^5 \cdot \sqrt{\sin x}}{(x^7-4)^2}\right)\\ =& \ln\Big((x^2+1)^5 \cdot \sqrt{\sin x}\Big) - \ln\big((x^7-4)^2\big)\\ =& \ln\big((x^2+1)^5\big) + \ln\Big(\sqrt{\sin x}\Big) - 2\ln\big(x^7-4\big)\\ =& 5\ln\big(x^2+1\big) + \frac{1}{2}\ln(\sin x) - 2\ln\big(x^7-4\big)\end{aligned}$$

Taking the derivative of both sides we get using the chain rule:

$$\frac{1}{y}\frac{dy}{dx} = \frac{5 \cdot 2x}{x^2+1} + \frac{\cos x}{2\sin x} - \frac{2 \cdot 6x^5}{x^7-4}$$

Multiplying both sides by y, and using that $y = f(x)$ we get that:

$$\frac{dy}{dx} = \left(\frac{(x^2+1)^5 \cdot \sqrt{\sin x}}{(x^7-4)^2}\right) \cdot \left(\frac{5 \cdot 2x}{x^2+1} + \frac{\cos x}{2\sin x} - \frac{2 \cdot 6x^5}{x^7-4}\right)$$

which is a very complicated expression, but was not that difficult to obtain.

3.4 Examples: More Physical Chain Rule and Implicit Differentiation

Example 3.16. We want to find the tangent line to:

$$e^{\cos(xy)} = 1$$

at $(\sqrt{\pi}/\sqrt{2}, \sqrt{\pi}/\sqrt{2})$. We take a derivative of both sides, and note that the derivative of the right hand side is obviously zero. For the right hand side we use the chain rule:

$$\begin{aligned}\frac{de^{\cos(xy)}}{dx} =& \frac{de^{\cos(xy)}}{d\cos(xy)} \cdot \frac{d\cos(xy)}{dxy} \cdot \frac{dxy}{dx}\\ =& e^{\cos(xy)}(-\sin(xy))(y + x\frac{dy}{dx})\end{aligned}$$

Equating the left hand side (the above) and the right hand side (which is zero), we get:

$$e^{\cos(xy)}(-\sin(xy))(y + x\frac{dy}{dx}) = 0$$

Since at $(\sqrt{\pi}/\sqrt{2}, \sqrt{\pi}/\sqrt{2})$, we know that $e^{\cos(xy)}(-\sin(xy))$ is non zero, we can divide both sides by it to get:

$$y + x\frac{dy}{dx} = 0$$

hence:

$$\frac{dy}{dx} = -\frac{y}{x}$$

so at $(\sqrt{\pi}/\sqrt{2}, \sqrt{\pi}/\sqrt{2})$ we have:

$$\frac{dy}{dx}(\sqrt{\pi}/\sqrt{2}, \sqrt{\pi}/\sqrt{2}) = -1$$

The tangent line is thus:

$$y - \frac{\sqrt{\pi}}{\sqrt{2}} = -\left(x - \frac{\sqrt{\pi}}{\sqrt{2}}\right)$$

Example 3.17. The force a proton exerts on an electron is given by:

$$F(r) = \frac{-k}{r^2}$$

where k is some constant, $F(r)$ is in Newtons, and r is in meters representing the distance between the electron and proton. The distance between the electron and proton is given by the function:

$$r(t) = 2 + \cos(\pi \cdot t)$$

We have that by the power rule:

$$\frac{dF}{dr} = \frac{2k}{r^3}$$

and that by the chain rule:

$$\frac{dr}{dt} = -\pi \cdot \sin(\pi t)$$

At $t = 1$, we have that:

$$\frac{dr}{dt}(1) = -\pi \cdot \sin(\pi) = 0$$

hence at $t = 1$:

$$\frac{dF}{dt}(1) = \frac{dF}{dr}(r(1)) \cdot \frac{dr}{dt}(1) = 0$$

Hence at $t = 1$ seconds we have that force the proton exerts on the electron is zero.

Example 3.18. The volume of a snowball is given by:

$$S(r) = \frac{4\pi r^3}{3}$$

where r is it's radius in meters. We have that:

$$r(t) = \frac{1}{(t+1)^2} - \frac{1}{12}$$

Taking a derivative of the volume with respect to time, we get that:

$$\begin{aligned}\frac{dS}{dt} =& \frac{dS}{dr} \cdot \frac{dr}{dt} \\ =& -8\pi r^2 \cdot \frac{1}{(t+1)^3}\end{aligned}$$

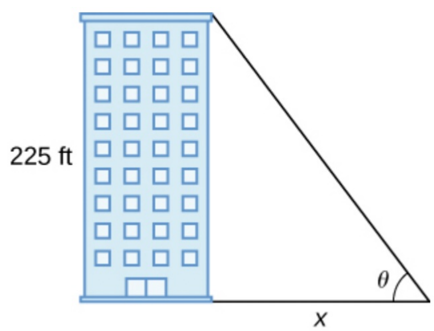
We want to know how quickly the volume is decreasing at $t = 1$ second. We know that:

$$\begin{aligned}r(1) =& \frac{1}{(1+1)^2} - \frac{1}{12} \\ =& \frac{1}{4} - \frac{1}{12} \\ =& \frac{1}{6}\end{aligned}$$

So:

$$\begin{aligned}\frac{dS}{dt}(1) =& -8\pi(1/6)^2 \cdot \frac{1}{2^3} \\ =& -\frac{\pi}{36}\end{aligned}$$

Example 3.19. In this example we solve problem 296 in the textbook. In particular, we have that a building casts a shadow of length x as the sun moves throughout the sky. We have the following diagram:



We want to find $d\theta/dx$ when $x = 272$ft. We know that

$$\tan\theta = \frac{225}{x}$$

hence we can apply take a derivative with respect to both sides to find that:

$$\sec^2\theta \cdot \frac{d\theta}{dx} = \frac{-225}{x^2}$$

hence:

$$\frac{d\theta}{dx} = \frac{-225}{x^2 \cdot \sec^2\theta}$$

We need to know what $\sec^2\theta$ is. Note that the hypotenuse of this triangle is:

$$\sqrt{x^2 + 225^2}$$

so:

$$\cos\theta = \frac{x}{\sqrt{x^2 + 225^2}}$$

so:

$$\cos^2\theta = \frac{x^2}{x^2 + 225^2}$$

hence:

$$\sec^2\theta = \frac{x^2 + 225^2}{x^2}$$

Plugging this into our formula for $d\theta/dx$ yields:

$$\frac{d\theta}{dx} = \frac{-225}{x^2 + 225^2}$$

plugging in $x = 227$ we obtain: that:

$$\frac{d\theta}{dx} \approx .002$$

We could alternatively say that:

$$\theta = \arctan\left(\frac{225}{x}\right)$$

hence:

$$\frac{d\theta}{dx} = \frac{1}{1 + (225/x)^2} \cdot \left(-\frac{225}{x^2}\right)$$

which gives the same answer when $x = 227$.

Week IV: Midterm Prep

Week V: Related Rates

5.1 Related Rates Problems I

Now that we know how to physically interpret derivatives, and in general how to take derivatives, it is time to begin seeing what derivative can tell us about physical situations. Our first class of applications are called related rates problems, and they rely heavily on the use of the chain rule. Often times, two (or more) quantities, say x and y can be related by a function; for example if you walking in a circle, and you position for all time is given by (x, y), then we know that if the radius of the circle is 2 meters, then we know that your x and y coordinates must satisfy:

$$x^2 + y^2 = 4$$

Since your x and y coordinates are changing with time, we can glean information about how fast we are moving in the x direction, and how fast we are moving in the y direction by taking a time derivative. In particular, since x is a function of time, we have that by the chain rule:

$$\frac{dx^2}{dt} = \frac{dx^2}{dx} \cdot \frac{dx}{t} = 2x \cdot \frac{dx}{dt}$$

and similarly that:

$$\frac{dy^2}{dt} = 2y\frac{dy}{dt}$$

We thus have that the velocity in the x direction, which is given by dx/dt, and the velocity in the y direction, which is given by dy/dt, must satisfy:

$$2x\frac{dx}{dt} + 2y\frac{dy}{dt} = 0$$

In particular, if we know our position, and the speed in either x direction, or the y direction, then we can figure out what the speed in the other direction. This is in essence the idea of related rates problems, and we will spend the next few lectures learning how to set them up, and solve them.

Let us begin with an abstract example:

Example 5.1. Suppose that x and y satisfy:

$$2x + y^3 = 10$$

If we know that $dx/dt = -3$ at $(1, 2)$ we want to determine what dy/dt is. We take a derivative of both sides and find that:

$$2 \cdot \frac{dx}{dt} + 3y^2 \cdot \frac{dy}{dt} = 0$$

We know that $dx/dt = -3$, $x = 1$ and $y = 2$, hence:

$$-6 + 3 \cdot 4 \cdot \frac{dy}{dt} = 0$$

Implying that $dy/dt = 1/2$.

Example 5.2. Suppose that x, y, and r satisfy:

$$x^2 + y^2 = r^2$$

In the context of our motivating example, this would be like walking in a helix, so walking in circles where the radius is continuously changing. The derivative satisfies:

$$2x\frac{dx}{dt} + 2y\frac{dy}{dt} = 2r \cdot \frac{dr}{dt}$$

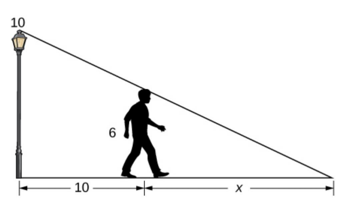
Now if $x = 6$, and $r = 10$, we know that by the pythagorean theorem $y = 8$. Moreover, if $dr/dt = 1$, $dy/dt = -3$, we can solve for dx/dt as follows:

$$6\frac{dx}{dt} - 24 = 10 \Rightarrow \frac{dx}{dt} = \frac{34}{6} \text{ m/s}$$

Bonus question, what do dx/dt, dy/dt, and dr/dt represent here?

Let's attempt a more fleshed out example:

Example 5.3. Lets look at the following picture:



In this problem, we have a 6ft man walking away from a 10ft tall lamp. As we walks, the tip of the shadow, gets further and further way from the lamp. Knowing that the man is walking at a constant speed of 3ft/s, we want to find how quickly the tip of his shadow is moving away from the pole. Let l be the distance between the poll and the man, and x be the distance between the man and his shadow. Since the triangles are similar, we know that:

$$\frac{6}{x} = \frac{10}{l+x}$$

Cross multiplying we find that:

$$6 \cdot (l+x) = 10x$$

Now we take a time derivative, and obtain that:

$$6\left(\frac{dl}{dt} + \frac{dx}{dt}\right) = 10 \cdot \frac{dx}{t}$$

Rearranging terms, we find that:

$$\frac{dx}{dt} = \frac{3}{2}\frac{dl}{dt}$$

We know that $dl/dt = 3$, $dx/dt = 9/2$ft/s. It follows that the speed at which the tip of the shadow is moving away from the man is $3 + 9/2 = 7.5$ft/s.

5.2 Related Rates Problems II

Example 5.4. Suppose that x and y depend on t, and satisfy the following equation:

$$xy + \sin(x+y) = 3$$

We want to find an equation relating dx/dt and dy/dt. Indeed, we have that by taking a derivative with respect to t of both sides:

$$\frac{d}{dt}(xy) + \frac{d}{dt}\sin(x+y) = 0$$

The product rule tells us that:

$$\frac{d}{dt}(xy) = \frac{dx}{dt}y + x\frac{dy}{dt}$$

while the chain rule tells us that:

$$\begin{aligned}\frac{d}{dt}\sin(x+y) &= \frac{d\sin(x+y)}{d(x+y)} \cdot \frac{d(x+y)}{dt} \\ &= \cos(x+y) \cdot \left(\frac{dx}{dt} + \frac{dy}{dt}\right)\end{aligned}$$

Putting everything together we get:

$$\frac{dx}{dt}y + x\frac{dy}{dt} + \cos(x+y) \cdot \left(\frac{dx}{dt} + \frac{dy}{dt}\right) = 0$$

Example 5.5. In class we discussed how to use the product rule to find the derivative of a product of more than two functions. In particular, if x, y, z, and w depend on t and if we have that the following equations relating them:

$$xyz = w^2$$

Taking a derivative with respect to t we get that:

$$\frac{d}{dt}(xyz) = \frac{d}{dt}(w^2)$$

For the right hand side we use the chain rule to get $2w\frac{dw}{dt}$. For the left hand side, we see that we can write xyz as $x \cdot (yz)$ (or $(xy) \cdot z$), and then use the product rule to get:

$$\begin{aligned}\frac{d}{dt}(xyz) =& \frac{d}{dt}(x \cdot (yz)) \\ =& \frac{dx}{dt} \cdot yz + x \cdot \frac{dyz}{dt}\end{aligned}$$

We will use the product rule again to get:

$$\begin{aligned}\frac{d}{dt}(xyz) =& \frac{dx}{dt} \cdot yz + x \cdot \left(\frac{dy}{dt}z + y\frac{dz}{dt}\right) \\ =& \frac{dx}{dt}yz + x\frac{dy}{dt}z + xy\frac{dz}{dt}\end{aligned}$$

Putting this all together we have that:

$$\frac{dx}{dt}yz + x\frac{dy}{dt}z + xy\frac{dz}{dt} = 2w\frac{dw}{dt}$$

If we have n functions $f_1, \ldots, f_n$ of x or t or whatever, then we get a generalized product rule:

$$\frac{d}{dx}(f_1 \cdots f_n) = \frac{df_1}{dx}f_2 \cdots f_n + f_1\frac{df_2}{dx}f_3 \cdots f_n + \cdots + f_1 \cdots \frac{df_{n-1}}{dx}f_n + f_1 \cdots f_{n-1}\frac{df_n}{dx}$$

Example 5.6. A 25-ft ladder is leaning against a wall. If we push the ladder toward the wall at a rate of 1 ft/sec, and the bottom of the ladder is initially 20 ft away from the wall, we want to find how fast the ladder is sliding up the wall after we have been pushing for 5 seconds. Let x the distance of the bottom of the ladder from the wall, and y the height of the top of ladder. If we have been pushing the ladder for 5 seconds, we have that $x = 15$, so:

$$15^2 + y^2 = 625 \Rightarrow y = \sqrt{625 - 225} = \sqrt{400} = 20$$

We know that:

$$x^2 + y^2 = 10$$

hence we have that:

$$2x\frac{dx}{dt} + 2y\frac{dy}{dt} = 0$$

We know that $x = 15$, $y = 20$, and $dx/dt = -5$ hence:

$$-5 \cdot 30 + 40\frac{dy}{dt} = 0 \Rightarrow \frac{dy}{dt} = 15/4\,\mathrm{ft}\,/\,\mathrm{s}$$

Example 5.7. Two airplanes are flying in the air at the same height: airplane A is flying east at 250 mi/h and airplane B is flying north at 300 mi/h. If they are both heading to the same airport, located 30 miles east of airplane A and 40 miles north of airplane B, we want to find the rate at which the distance is changing. Let Δ be the distance between the two planes, x be the distance of airplane A from the airport, and y be the distance of airplane b from the airport. We have that:

$$\frac{dx}{dt} = -250 \qquad \text{and} \qquad \frac{dy}{dt} = -300$$

The distance is related to x and y by the equation:

$$\Delta^2 = x^2 + y^2$$

hence:

$$\Delta \frac{d\Delta}{dt} = x\frac{dx}{dt} + y\frac{dy}{dt}$$

We have that $x = 30$, $y = 40$, and so $\Delta = \sqrt{30^2 + 40^2} = 50$mi. We can thus see that:

$$\frac{d\Delta}{dt} = \frac{-30 \cdot 250 - 300 \cdot 40}{50} = -30 \cdot 5 - 6 \cdot 40 = -150 - 240 = -390$$

do the distance between the two airplanes is shrinking at a speed of 390 miles per hour.

Example 5.8. Two buses are driving along parallel freeways that are 5 mi apart, one heading east and the other heading west. Assuming that each bus drives a constant 55 mph, we want to find the rate at which the distance between the buses is changing when they are 13 mi apart, heading away from each other.

First note that the distance Δ between the two busses is given by:

$$\Delta^2 = x^2 + 5^2$$

where x is how far they are apart on the highway, and 5 is how far apart the highways are. Taking a time derivative we have that:

$$2\Delta\frac{d\Delta}{dt} = 2x\frac{dx}{dt}$$

Note that $dx/dt = 110$ as *each bus* is moving at 55 miles per hour away from each other. When they are 13 miles apart, we have that:

$$13^2 = x^2 + 5^2 \Rightarrow x^2 = 144 \Rightarrow x = 12$$

hence:

$$\frac{d\Delta}{dt} = \frac{12}{13} \cdot 110 \approx 101.5\,\text{mph}$$

5.3 Related Rates Problems III

We continue with examples of related rates problems. First the solution to todays daily warm up:

Example 5.9. Suppose that x, y, z and w depend on t, and satisfy:

$$x^2y^2z^2 = \cos(w)$$

We take a derivative with respect to t of both sides. The chain rule tells us that the right hand side is given by:

$$\frac{d}{dt}\cos(w) = -\sin(w)\frac{dw}{dt}$$

while the product rule tells us that the left hand side is given by:

$$\begin{aligned}\frac{d}{dt}(x^2y^2z^2) =& 2x\frac{dx}{dt}y^2z^2 + x^2 \cdot \frac{d}{dt}(y^2z^2)\\ =& 2x\frac{dx}{dt}y^2z^2 + x^2 \cdot \left(2y\frac{dy}{dt} + 2y^2z\frac{dz}{dt}\right)\\ =& 2xy^2z^2\frac{dx}{dt} + 2yx^2z^2\frac{dy}{dt} + 2zx^2y^2\frac{dz}{dt}\end{aligned}$$

It follows that:

$$2xy^2z^2\frac{dx}{dt} + 2yx^2z^2\frac{dy}{dt} + 2zx^2y^2\frac{dz}{dt} = -\sin(w)\frac{dw}{dt}$$

Example 5.10. The radius of a sphere is increasing at a rate of 9 cm/ sec. We want to find the radius of the sphere when the volume and the radius of the sphere are increasing at the same numerical rate. We have that:

$$V = \frac{4}{3}\pi r^3$$

We thus have that:

$$\frac{dV}{dt} = 4\pi r^2 \frac{dr}{dt}$$

We have to set these equal to each other:

$$4\pi r^2 \frac{dr}{dt} = \frac{dr}{dt}$$

Solving for r we find that:

$$r = \frac{1}{2\sqrt{\pi}}$$

Example 5.11. The base of a triangle is shrinking at a rate of 1 cm/min and the height of the triangle is increasing at a rate of 5 cm/ min. We want to ind the rate at which the area of the triangle changes when the height is 22 cm and the base is 10 cm.

Recall that the area of a triangle is $1/2b \cdot h$. It follows that:

$$\frac{dA}{dt} = \frac{1}{2}\left(\frac{db}{dt}h + b\frac{dh}{dt}\right)$$

Plugging everything we know into the above equation:

$$\frac{dA}{dt} = \frac{1}{2}(-22 + 10 \cdot 5) = 14\,\mathrm{cm}\,/\,\mathrm{min}$$

Example 5.12. In this example we consider a right conical tank of radius 5m and height 16m. We want to know how fast the depth of the water is changing when the water is 10m high, and water leaks at a rate of $10\,\mathrm{m}^3$.

Using similar triangles we have that if r and h are the radius and height of the cone that the current volume of water takes up then:

$$\frac{5}{16} = \frac{r}{h}$$

so:

$$r = \frac{5}{16}h$$

Thus the volume of the cone as a function of h is given by:

$$V = \frac{\pi}{3}\left(\frac{5}{16}\right)^2 \cdot h^3$$

Taking a derivative we obtain that:

$$\frac{dV}{dt} = \pi \cdot \left(\frac{5}{16}\right)^2 h^2 \cdot \frac{dh}{dt}$$

We know that dV/dt is -10, and we know that $h = 10$, hence:

$$\frac{dh}{dt} = -\frac{1}{10 \cdot \pi \left(\frac{5}{16}\right)^2}\,\mathrm{m}\,/\,\mathrm{s}$$

Week VI: Optimization

6.1 Local Maxima and Minima

Before discussing maxima and minima we review the daily warm up problem:

Example 6.1. A spherical balloon is inflating at a constant rate of $2\,\mathrm{cm}^3\,/\,\mathrm{s}$. We want to find dr/dt when $V = 36\pi\,\mathrm{cm}^3$. We have that:

$$V = \frac{4}{3}\pi r^3$$

Taking a derivative with respect to time and making the substitution $dv/dt = 2$ we obtain that:

$$2 = 4\pi r^2 \frac{dr}{dt}$$

All we need to do is find r. We have that:

$$36\pi = \frac{4}{3}\pi r^3 \Rightarrow 36 = \frac{4}{3}r^3$$

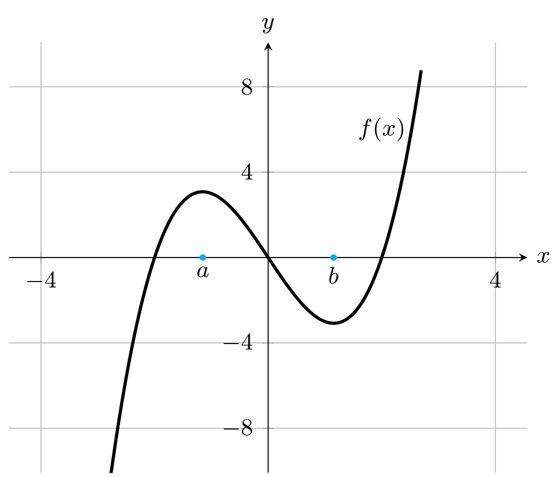
Multiplying both sides by $3/4$ we get that:

$$r^3 = 3 \cdot 9 = 27 \Rightarrow r = 3$$

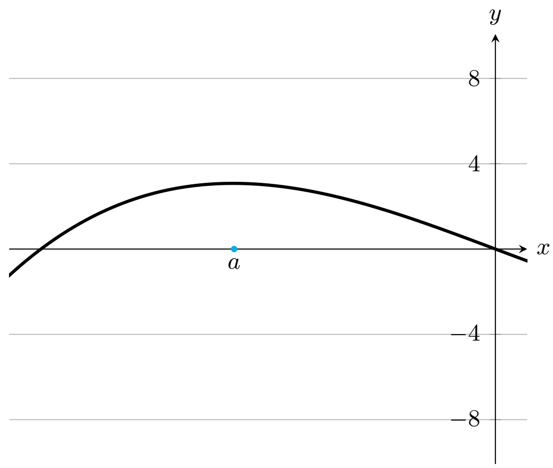
It follows that:

$$\frac{dr}{dt} = \frac{1}{18\pi}\,\mathrm{cm}\,/\,\mathrm{s}$$

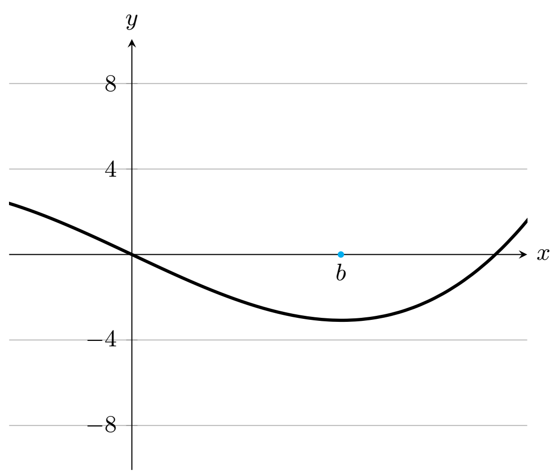
To discuss minima and maxima, we must first interpret what it means for the derivative of a function to be zero, or not exist at a point. Consider the following graph:



Clearly, something is happening at the points a and b. If we were to zoom in around a it would appear that the maximum value of f occurs at a:

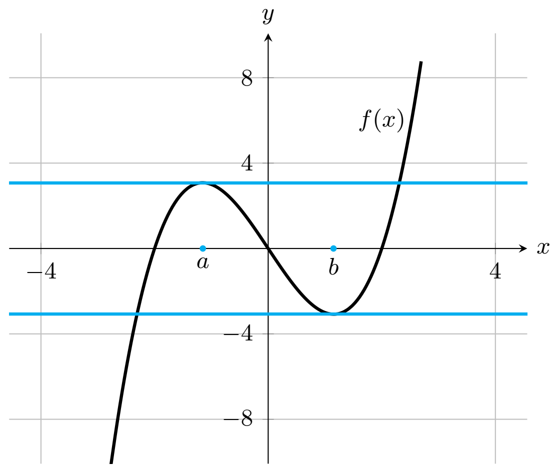


and if we zoom in around b it would look like the minimum value of f occurs at b:



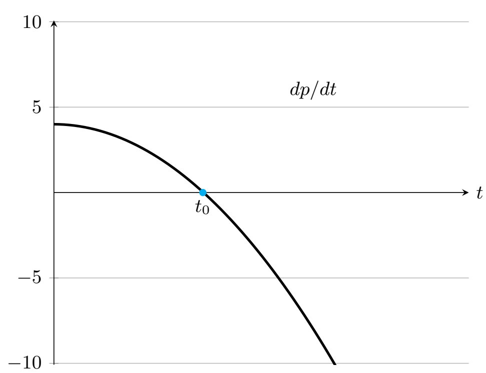
The points a and b are what call *local extrema*, with a being a *local maximum* and b being a *local minimum*. The use of the word local here just means that if we zoom in around a or b they it looks like the maximum/minimum value occurs at a/b. How do we characterize this mathematically?

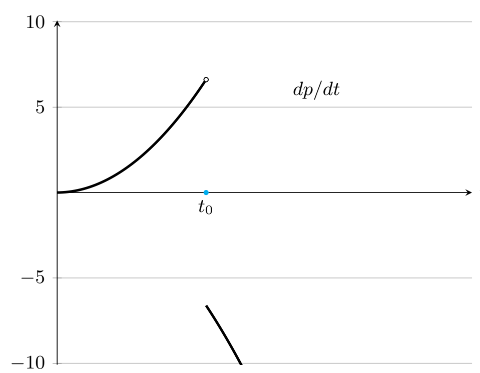
Well, we see that when a differentiable function obtains a local maximum or minimum the value of it's derivative is zero at that point. Indeed, we can draw the tangent lines of f at a and b and clearly see that the slope is zero:



So we know that the place to look for local minimums and local maximums is where the derivative of the function is zero. How do we tell whether a local extrema is a minimum or a maximum though? Well before a we have that the derivative is positive so the function is increasing, and between a and b the derivative is negative so the function is decreasing. If something is increasing until a and then is decreasing after a then it must be the case that the local maximum occurs at a. Similarly, after b the derivative is positive, so the function is increasing, hence before b the function is decreasing, and after b the function is increasing hence b must be a local minimum.

To further drive home this point we consider a real life example. If $p(t)$ is the population of rabbits in a field as a function of time, we know that $p(t)$ is increasing at some points and decreasing at others. In particular, in the spring the population of rabbits should be increasing, and in the winter the population of rabbits in the field will be decreasing as the food supply decreases, and the old rabbits die. In particular, in the spring months the derivative of $p(t)$ is positive, and the winter months the derivative of $p(t)$ is negative, so at some point, t_0, in between the winter and spring months the derivative has to be zero, or not exist. Intuitively, the only way the derivative of a function can go from positive to negative is if it crosses over the x-axis and so is zero, or if it doesn't exist at that point and so before that point is positive, and after that point is negative. In other words we have the following two possible pictures for the derivative of $p(t)$:





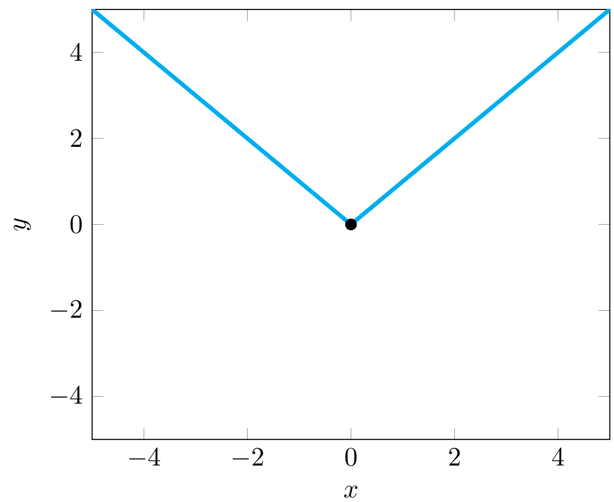
So either the derivative of p becomes zero to change sign, or hops over it to change sign. In either case, we have that the population obtains its maximum where the derivative is either zero or does not exist because this is precisely where the function stops growing. With all the above in mind we have the following definition:

Definition 6.1. Let f be a function then a real number a is a **critical point of** f if $\frac{df}{dx}(a)$ equals zero or does not exist. A critical point is a **local minimum** if there is an interval of the form $[b, a]$ such

df/dx is negative, and an interval of the form $[a,c]$ such that df/dx is greater than or equal to zero. A critical point is a **local maximum** if there is an interval of the form $[b,a]$ such df/dx is greater than or equal to zero, and an interval of the form $[a,c]$ such that df/dx is less than or equal to zero.

Note that any such interval can be taken to have end points to be critical points because the only place where the derivative can switch sign is at a critical point. We look at two examples:

Example 6.2. Suppose that $f(x) = |x|$, then we have the following graph:



Clearly at $x = 0$, $|x|$ obtains a minimal value of zero at $x = 0$. We can think of $|x|$ as the pieceiwse function:

$$|x| = \begin{cases} -x & \text{for } x \leq 0 \\ x & \text{for } x > 0 \end{cases}$$

The derivative is then:

$$\frac{d}{dx}(|x|) = \begin{cases} -1 & \text{for } x < 0 \\ 1 & \text{for } x > 0 \end{cases}$$

and does not exist at 0.

Example 6.3. Consider:

$$f(x) = x^3 - 27x$$

We want to find the critical points of this function. Taking a derivative of both sides:

$$\frac{df}{dx} = 3x^2 - 27$$

We want to set this equal to zero to find the critical points:

$$3x^2 - 27 = 0 \Rightarrow x^2 = 9 \Rightarrow x = \pm 3$$

The only critical points are 3 and -3, so we can test any number less than -3, in between -3 and 3, and greater than three, and the sign of the derivative will be fixed. In particular, since $(-4^2) - 9 = 7$ which is positive we know that the sign of the derivative is positive for all numbers in $(-\infty, 3)$. Moreover, since $0^2 - 9 = -9 < 0$ we know that for all numbers in the interval $(-3, 3)$ the derivative is negative. Finally since $4^2 - 9 = 7$ ich is positive we know that the derivative is positive for all numbers in $(3, \infty)$. Putting it all together -3 is a local maximum because the function goes from increasing to decreasing at -3, and 3 is a local minimum because the function goes from decreasing to increasing 3.

6.2 Global Minima and Maxima

We go over the daily warm up:

Example 6.4. Consider:

$$f(x) = \frac{x^3}{3} - 16x$$

We want to find the critical points of this function. Taking a derivative of both sides:

$$\frac{df}{dx} = x^2 - 16$$

We want to set this equal to zero to find the critical points:

$$x^2 - 16 = 0 \Rightarrow x^2 = 4 \Rightarrow x = \pm 4$$

The only critical points are 4 and -4, so we can test any number less than -4, in between -4 and 4, and greater than three, and the sign of the derivative will be fixed. In particular, since $(-5^2)-16 = 9$ which is positive we know that the sign of the derivative is positive for all numbers in $(-\infty, 4)$. Moreover, since $0^2 - 16 = -16 < 0$ we know that for all numbers in the interval $(-4, 4)$ the derivative is negative. Finally since $5^2 - 16 = 9$ ich is positive we know that the derivative is positive for all numbers in $(4, \infty)$. Putting it all together -4 is a local maximum because the function goes from increasing to decreasing at -4, and 4 is a local minimum because the function goes from decreasing to increasing 4.

Before discussing global minima and maxima, we look at one important example of finding local minima and maxima.

Example 6.5. Consider the function:

$$f(x) = \sin x$$

the derivative of $\sin x$ is $\cos x$, hence the critical points occur when $\cos x = 0$. It follows that we have a critical point at $\pi/2 + n\pi$ for every integer n. We claim that if n is even then $\pi/2 + n\pi$ is local maximum and if n is odd then $\pi/2+n\pi$ is a local minimum. Note that this should make sense because if n is even then $n = 2m$ for some number m so:

$$\sin(\pi/2 + 2m\pi) = \sin(\pi/2) = 1$$

which is the maximum value of $\sin x$. If n is odd then $n = 2m + 1$ so:

$$\sin(\pi/2 + (2m_1)\pi) = \sin(\pi/2 + \pi + 2m\pi) = \sin(3\pi/2) = -1$$

which is the minimum value of $\sin x$.

To actually argue this we know that $\pi/2 + 2m\pi$ describes an angle at the top of the unit circle, i.e. it describes the same angle as $\pi/2$. In particular, if we chose an angle slightly less than $\pi/2$ then cos of this angle is positive, and we chose an angle slightly more than $\pi/2$ then cos of this angle is negative. It follows that $\pi/2 + n\pi$ is a maximum if n is even.

If n is odd then we have that:

$$\pi/2 + (2m + 1)\pi = 3\pi/2 + 2m\pi$$

so this defines an angle at the bottom of the unit circle, i.e. it describes the same angle as $3\pi/2$. In particular, if we chose an angle slightly less than $3\pi/2$ then cos of this angle is negative, and we chose an angle slightly more than $3\pi/2$ then cos of this angle is positive. It follows that $\pi/2 + n\pi$ is a local minimum if n is odd.

Recall in Example 6.3 we found the local minima and maxima of:

$$f(x) = x^3 - 27x$$

These were indeed extremely local, as the function $f(x)$ does not attain "global maximum" or a "global minimum". In other words, there is no real number a such that $f(a)$ is bigger or smaller than any other value of $f(x)$. This is because after the critical point 3 $f(x)$ continues to grow and just gets bigger, and before -3, f has been growing from $-\infty$. In particular, for whatever number you write down, I can write down a number equal to $f(a)$ for some a which is bigger or smaller than your number.

The thing to notice is that if we instead define our function on a closed interval $[b, a]$ instead of all of the real numbers, then we do attain a global maximum or minimum! Indeed, since the function can't grow or decrease for ever anymore, we will have that a global maximum or minimum will always occurs at the end points of a function or the critical points. This is such an important idea we upgrade it to a theorem:

Theorem 6.1. *Let f be a function defined on a closed interval $[b, a]$. Then f attains a global maximum and a global minimum on f which occurs either at a, b, or one of the critical points of f.*

We have the following example:

Example 6.6. Consider the function

$$f(x) = x^3 - 27x$$

again. As we saw this function had a local maximum at -3 and a local minimum at $+3$. We restrict the function $[-4, 4]$, and we calculate:

$$f(-4) = 44 \qquad f(-3) = 54 \qquad f(3) = -54 \qquad f(4) = -44$$

It follows that on this interval -3 is a maximum, and 3 is a minimum.

If instead we chose the interval $[-7, 7]$ we have that $f(-7) = -154$ and $f(7) = 154$, hence -7 and 7 are the global maximum, and global minimum.

In some cases we can find global minimums or maximums of functions that are defined for all real numbers. These situations are rarer, and occur when you declare with confidence that the function does not get any larger (respectively smaller/more negative) than the global maximum (respectively the global minimum). As such a case consider the following example:

Example 6.7. Let:

$$f(x) = x^2$$

then we have a critical point at 0 which isa minimum because for $x < 0$ we have that $2x$ is negative, and for $x > 0$ we have that $2x$ is positive. Since the function is always decreasing before 0 and always increasing after 0 we can safely say that 0 is a global minimum. Moreover, we can also argue that 0 is the global minimum because squaring a number always gives you something greater than or equal to zero.

6.3 Examples

We begin with the daily warmup:

Example 6.8. We consider the function $f(x) = x^2 + 2x - 1$ and want to first find it's critical points, and determine whether they are local minimums or local maximums. We first take a derivative:

$$\frac{df}{dx} = 2x + 2$$

We know that this derivative exists everywhere so we need only check where it is equal to zero:

$$2x + 2 = 0 \Rightarrow 2x = -2 \Rightarrow x = -1$$

For $x < -1$ we can plug in -2 and find that $\frac{df}{dx}(-2) = -2$ hence f decreasing before -1. Similarly, $\frac{df}{dx}(2) = 2$, and so the f is increasing after -1. It follows that -1 is a local minimum.

We now want to find the global minimum/maximum of $x^2 + 2x - 1$ on the interval $[-2, 3]$. First note that since -1 is the only critical point of f, and it is a local minimum, that is also a global minimum on any interval containing -1. This is because f is decreasing $(-\infty, 0)$ and increasing on $(0, \infty)$, hence the function will never be smaller than it's value at -1. When in doubt though, we can always directly compute the value of f at the critical points and end points:

$$f(-1) = -2 \qquad \text{and} \qquad f(-2) = -1 \qquad \text{and} \qquad f(3) = 14$$

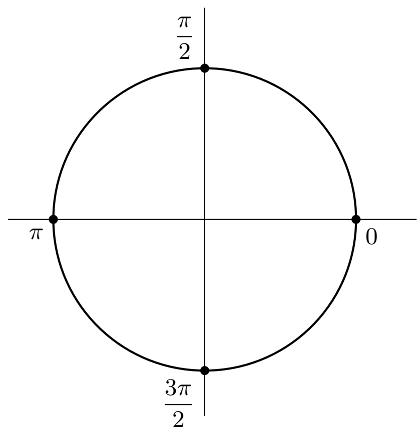
so $x = -1$ is the global minimum, and $x = 3$ is the global maximum.

We now turn to a more complicated example:

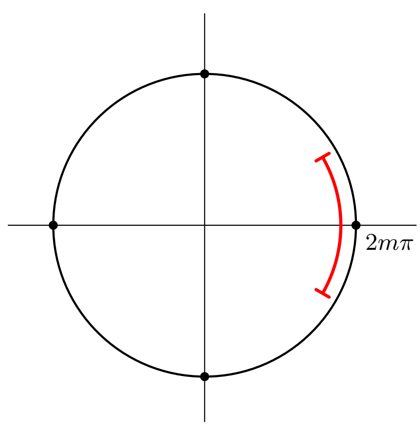
Example 6.9. We want to find all critical points of $\cos\theta$. Indeed, we know that:

$$\frac{d}{d\theta}(\cos\theta) = -\sin\theta$$

This derivative exists everywhere so we need only determine when $-\sin\theta = 0$. Looking at our unit circle:

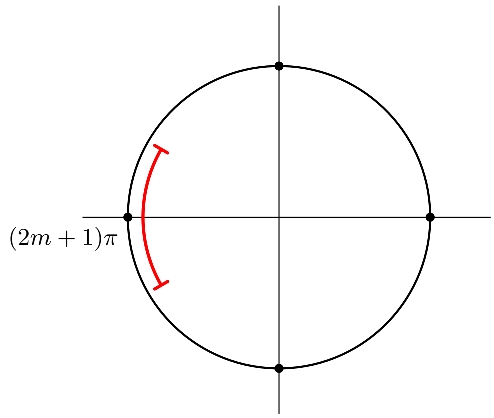


we see that since $\cos\theta$ is the x coordinate, $\sin\theta$ is the y coordinate, that $-\sin x$ is only zero when $\theta = n\pi$ where n is a whole number, positive or negative. We will determine when a critical point is a local minimum or maximum. We first consider the case where n is even, i.e. $n = 2 \cdot m$ for some integer m. In particular, we $x = 2m\pi$ then we know we are at the point $(1, 0)$ on the unit circle. Consider the following picture:



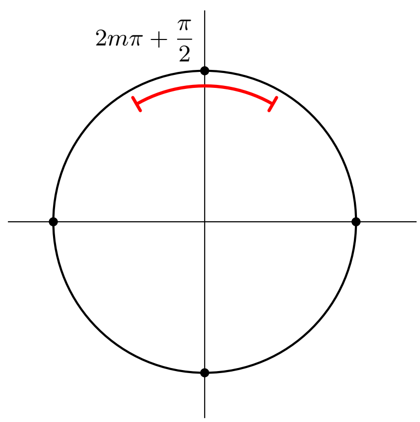
From the picture we see that if for an angle less than $2m\pi$ we have that $\sin x$ is negative hence $-\sin\theta$ is positive, and for an angle greater than $2m\pi$ we have that $\sin\theta$ is positive. It follows that before $2m\pi$ $\cos\theta$ is increasing, and after $2m\pi$ $\cos\theta$ is decreasing so $2m\pi$ is a local maximum. Therefore, for even integers greater than zero we have that $n\pi$ is a local maximum.

If we consider and odd integer, then we have that $n = 2m + 1$ for some integer m. In particular, for $x = (2m + 1)\pi$ we are at the point $(-1, 0)$ on the unit circle. Examine the following picture:



From the picture we see that for an angle greater than $(2m+1)\pi$ $\sin\theta$ is negative so $-\sin\theta$ is positive, and for an angle less than $(2m + 1)\pi$ we have that $\sin\theta$ is positive so $-\sin\theta$ is negative. It follows that $(2m + 1)\pi$, $\cos\theta$ goes from decreasing to increasing and so $(2m + 1)\pi$ is a local minimum. In particular, for all odd integers, we have that $n\pi$ is a local minimum.

Example 6.10. We now do the same thing with $\sin\theta$. The derivative of $\sin\theta$ is $\cos\theta$ which exists everywhere and is zero when $\theta = n\pi + \frac{\pi}{2}$ for any integer n. We play the same game as we did in the previous example. First assume n is greater than equal to zero; n is either even or odd, so we suppose that n is even. If n is even and positive, then we have that $n = 2m$ for some positive integer m, and so the angle is of the form $2m\pi + \frac{\pi}{2}$. This puts us at the top of the unit circle circle, i.e. at $(0, 1)$ We have the following picture:

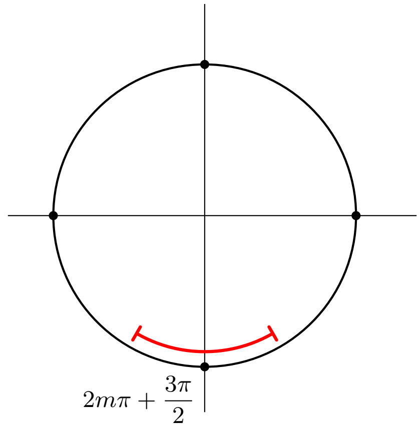


For θ less than $2m + \pi/2$ we have that $\cos\theta$ is positive, and for θ greater than $2m + \pi/2$ we have that $\cos\theta$ is negative. It follows that $\sin\theta$ goes from increasing to decreasing at $2m + \pi/2$, and is thus a local maximum. It follows that for all even integers, $n\pi + \pi/2$ is a local maximum.

Now we consider n an odd integer, i.e. $n = 2m + 1$ for some integer m. In this situation, we have that:

$$n\pi + \frac{\pi}{2} = (2m + 1)\pi + \frac{\pi}{2} = 2m\pi + \frac{3\pi}{2}$$

It follows that we are at the bottom of the unit circle and thus have the following picture:



We leave it as an exercise to the reader to argue that this point is local minimum, and thus for odd integers n, $n\pi + \pi/2$ is a local minimum.

Example 6.11. Suppose we have two positive numbers x and y such that $x + y = 10$. We want to answer the following question: what is the largest possible product $x \cdot y$? It turns out with optimization techniques we can answer such questions. Indeed, since $x + y = 10$, we know that $y = 10 - x$ hence the product $x \cdot y$ is given by $p(x) = x(10 - x) = 10x - x^2$. Moreover, we know that x has to be greater than or equal to zero and less than or equal to 10, hence we need to maximize the function $x(10 - x)$ on the interval $[0, 10]$. The endpoints obviously can't be maxima as at 0 and 10 the product of x and y is zero. We instead take a derivative of $p(x)$ and set it equal to zero:

$$\frac{dp}{dx} = 10 - 2x = 0$$

this implies that $x = 5$. If we plug in $x = 4$ we find that dp/dx is positive, and we plug in $x = 6$ we find that dp/dx is negative hence $x = 5$ is a maximum. It follows that this a global maximum as we are on a closed interval, and moreover that $y = 10 - 5 = 5$. The largest possible product of x and y is thus 25.

6.4 Optimization Examples

We start with the daily warm up:

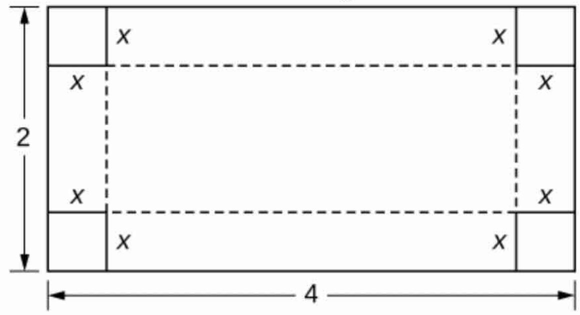
Example 6.12. We had $f(x) = x^3 - 6x - 1$, and we first wanted to find the critical points, and determine whether they are minima or maxima. The derivative of this function exists everywhere, and so we have to set its derivative equal to zero:

$$\frac{df}{dx} = 3x^2 - 6 = 0$$

This implies that $x = \pm\sqrt{2}$ are the critical points. We plug in -2 and we get $12 - 6 = 6$, so for $x < -\sqrt{2}$ we have that df/dx is positive. We plug in 0 and get $0 - 6 = -6$ and so for $-\sqrt{2} < x < \sqrt{2}$, df/dx is negative. We plug in 2 and get that df/dx is positive for $x > \sqrt{2}$. It follows that the function goes from increasing to decreasing at $-\sqrt{2}$ and from decreasing to increasing at $\sqrt{2}$. We therefore have that $-\sqrt{2}$ is a local minimum, and $\sqrt{2}$ is a local maximum.

Example 6.13. We solve the following problem from the textbook:

316. You are constructing a cardboard box with the dimensions 2 m by 4 m. You then cut equal-size squares from each corner so you may fold the edges. What are the dimensions of the box with the largest volume?



Note that x has to be greater than or equal to zero, and less than or equal to 1, as other wise we take away more width of the rectangle than exists. The heigh of the box will always be x, the length will be $4-2x$, and the width will be $2-2x$. It follows that:

$$\begin{aligned} V(x) =& x(4-2x)(2-2x) \\ =& 4x(2-x)(1-x) \\ =& 4x(x^2-3x+2) \\ =& 4(x^3-3x^2+2) \end{aligned}$$

We want to maximize this function on the interval $[0,1]$. We take a derivative and set it equal to zero:

$$\frac{dV}{dx} = 4(3x^2-6x+2) = 0$$

We divide both sides by 4 to obtain that:

$$3x^2-6x+2=0$$

We need to solve this quadratic, and we can either do it by using the quadratic equation:

$$\begin{aligned} x =& \frac{-b \pm \sqrt{b^2-4ac}}{2a} \\ =& \frac{6 \pm \sqrt{36-24}}{6} \\ =& 1 \pm \frac{\sqrt{12}}{6} \\ =& 1 \pm \frac{\sqrt{3}}{3} \\ =& 1 \pm \frac{1}{\sqrt{3}} \end{aligned}$$

Or we an complete the square by first subtracting 2 from both sides:

$$3x^2-6x=-2$$

Dividing by 3 throughout:

$$x^2-2x=-\frac{2}{3}$$

adding $(b/2)^2 = (2/2)^2 = 1$ to both sides:

$$x^2 - 2x + 1 = 1 - \frac{2}{3}$$

Noting that $x^2 - 2x + 1 = (x-1)^2$ hence:

$$(x-1)^2 = \frac{1}{3}$$

and then taking the $\pm$ square root of both sides, and adding one to obtain that:

$$x = 1 \pm \frac{1}{\sqrt{3}}$$

Regardless we know that $V(x)$ has critical points at $1 \pm \frac{1}{\sqrt{3}}$, however $1+1/\sqrt{3}$ is outside of our domain hence we needn't consider it. It suffices to confirm that $1-1/\sqrt{3}$ is a maximum. If we plug in 0 to the derivative we get $+2$ hence before $1-1/\sqrt{3}$ volume is increasing. If we plug in 1 we get -1 so after $1-1/\sqrt{3}$ volume is increasing. It follows that $1-1/\sqrt{3}$ is a global maximum, and thus the largest possible volume is given by:

$$\begin{aligned} V(1-1/\sqrt{3}) =& 4 \cdot (1-\sqrt{3})(1+1/\sqrt{3})(1/\sqrt{3}) \\ =& 4 \cdot (1-1/3)(1\sqrt{3}) \\ =& \frac{8}{3\sqrt{3}} \text{ m}^3 \end{aligned}$$

Week VII: The Second Derivative

6.1 One More Day on Optimization Examples

Todays daily warm up:

Example 6.1. The first question asked us to determine the difference between local extrema, and global extrema. The difference is is that local extrema look like peaks or valleys, i.e. if zoom in around a local minimum of maximum, it looks like like the top of a hill or the bottom of a valley. Global extrema are just where the function attains it's smallest or largest value. Local minima and local maxima can be global minima and maxima and vice versa, but this need not be the case.

We were tasked with find the global extrema of the function $\sin x$ on the interval $[-\pi/4, 3\pi/4]$. The derivative of $\sin x$ is $\cos x$ which exists everywhere, hence to find critical points we need only check where $\cos x$ is zero in the interval $[-\pi/4, 3\pi/4]$. This only happens when $x = \pi/2$, which is a local maximum because for $\cos x$ goes from positive to negative at $\pi/2$, as is easily seen by drawing a unit circle.

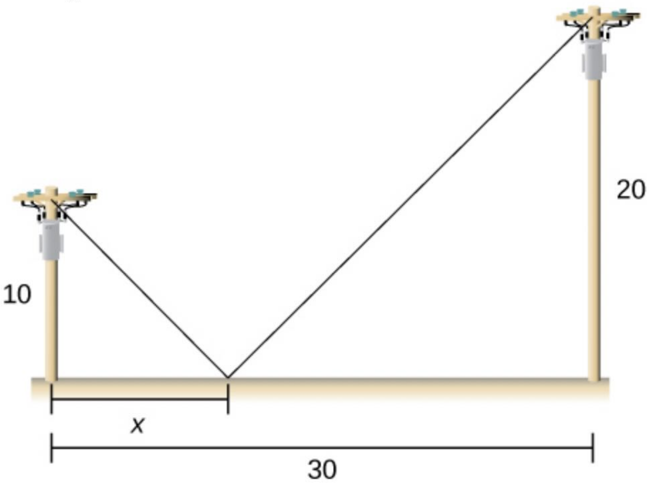
We now test our end points and critical points:

$$\sin(-\pi/4) = -\frac{\sqrt{2}}{2} \qquad \sin(\pi/2) = 1 \qquad \sin(3\pi/4) = \frac{\sqrt{2}}{2}$$

It follows that $-\pi/4$ is a global minimum, and $\pi/2$ is a global maximum.

Example 6.2. We do problem 322 from the textbook:

322. Two poles are connected by a wire that is also connected to the ground. The first pole is $20\,\text{ft}$ tall and the second pole is $10\,\text{ft}$ tall. There is a distance of $30\,\text{ft}$ between the two poles. Where should the wire be anchored to the ground to minimize the amount of wire needed?



Let l_1 denote the length of the wire coming from the left pole, and l_2 denote the length of the wire coming from the white pole. We want to minimize $l_1 + l_2$. Note that:

$$l_1 = \sqrt{10^2 + x^2} \qquad \text{and} \qquad l_2 = \sqrt{20^2 + (30 - x)^2}$$

and that $0 \le x \le 30$. We have that:

$$\frac{dl_1}{dx} = \frac{x}{\sqrt{10^2 + x^2}} \qquad \text{and} \qquad \frac{dl_2}{dx} = \frac{-(30 - x)}{\sqrt{20^2 + (30 - x)^2}}$$

We want to find x such that:

$$\frac{dl_1}{dx} + \frac{dl_2}{dx} = 0 \Rightarrow \frac{dl_1}{dx} = -\frac{dl_2}{dx}$$

hence we need to solve:

$$\frac{x}{\sqrt{10^2 + x^2}} = \frac{30 - x}{\sqrt{20^2 + (30 - x)^2}}$$

cross multiplying, and squaring both sides yields:

$$x^2(20^2 + (30 - x)^2) = (30 - x)^2(10^2 + x^2)$$

Expanding both sides gives:

$$400x^2 + x^2(30 - x)^2 = 100(30 - x)^2 + x^2(30 - x^2)$$

hence we need to solve :

$$400x^2 - 100(30 - x)^2 = 0$$

Dividing by 100 gives:

$$4x^2 - (30 - x)^2 = 0$$

Expanding $(30-x^2) = x^2 + 900 - 60x$ gives:

$$3x^2 + 60x - 900 = 0$$

Dividing by 3 yields:

$$x^2 + 20x - 300 = 0$$

We can factor $x^2 + 20x - 300$ as $(x+30)(x-10)$, hence $x = -30$ or $x = 10$. The first point is outside of our allowed x values, so we throw it out. For $x = 10$, we see that this a local minimum as:

$$\frac{dl_1}{dx}(0) + \frac{dl_2}{dx}(0) = 0 - \frac{30}{\sqrt{20^2 + 30^2}} < 0$$

while:

$$\frac{dl_1}{dx}(30) + \frac{dl_2}{dx}(30) = \frac{30}{\sqrt{20^2 + 30^2}} > 0$$

hence before $x = 10$ we have that $l_1 + l_2$ is decreasing, and after $x = 10$, $l_1 + l + 2$ is increasing. It follows that since this is only local minimum on the whole interval, it must be a global minimum, hence the minimum amount of wire needed is:

$$l_1(10) + l_2(10) = \sqrt{2 \cdot 10^2} + \sqrt{2 \cdot 20^2} = 30 \cdot \sqrt{2}$$

Example 6.3. Let x and y be positive integers such that $x + y = 10$. We want to minimize and maximize:

$$f(x,y) = y - \frac{1}{x}$$

Note that no such minimum exists, as as x gets closer to zero, $y - 1/x$ goes to negative infinity. We can find a global maximum on the interval $0 \leq x \leq 10$ though. Indeed, $y = 10 - x$, hence:

$$\frac{df}{dx} = -1 + \frac{1}{x^2} = 0$$

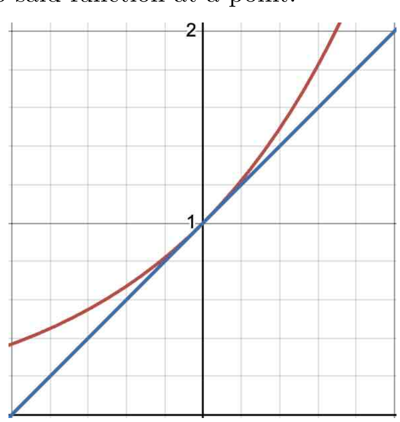
which is equal to zero if and only if $x = \pm 1$. We can't have $x = -1$ though, so we conclude that $x = 1$. If we plug in $x = 1/2$ we clearly get that $df/dx > 0$, while if we plug in 2 we obtain $df/dx < 0$. It follows that $x = 1$ is a global max, and thus the maximum value is:

$$10 - 1 - \frac{1}{1} = 8$$

6.2 Interpreting the Second Derivative

Recall that we first introduced the derivative in the context of finding the instantaneous rate of change of a function at a point. In particular, we argued that if $x(t)$ we a function describing the distance some object had traveled in t seconds, then dx/dt would be the velocity function of that object, telling us how fast the object is moving at t seconds.

We then further argued that mathematically, in the absence of physical the derivative should tell us the slope of the tangent line to said function at a point:



We also claimed that the tangent line offered us a good approximation for the function close to that point. In other words, as we zoom in, the tangent line and the function become indistinguishable from each other. Moreover, as we have seen in prior sections, the derivative tells us whether the function is increasing, or decreasing, or staying the same at any point.

The question we seek to answer this section, is what does the second derivative tell us, both mathematically, and physically? First note that by the second derivative we mean the function one obtains after taking a derivative twice. For example, we have that the first derivative of $\sin x$ is $\cos x$, so the second derivative of $\sin x$ is the derivative of $\cos x$ which is $-\sin x$. One can do this for any function who's derivative is differentiable.

We begin by arguing what the second derivative should be physically, based of what we know about a derivative. Consider the following example:

Example 6.4. Let $x(t) = 3t^2 + 2t + 1$ describe the distance of a particle from the origin for $t \geq 0$. The first derivative tell us the velocity function:

$$v(t) = \frac{dx}{dt} = 6t + 2$$

What should the second derivative tell us? Well $d^2x/dt^2 = dv/dt$, and a velocity function has units $\mathrm{m}\,/\,\mathrm{s}$. We know that the derivative should reflect a rate of change, so what is the rate of change of velocity? It is acceleration, which has units:

$$\frac{\mathrm{m}\,/\,\mathrm{s}}{\mathrm{s}} = \mathrm{m}\,/\,\mathrm{s}^2$$

With words we say acceleration has units 'meters per second squared'. In particular, the sign of the derivative tells us whether distance from the origin is increasing or decreasing, while the sign of the second derivative tells us whether velocity is increasing or decreasing.

Example 6.5. If $P(t)$ tells us the population of rabbits in a meadow over t months, then dP/dt tells us how the population is changing at any instance in time. It has units of rabbits per month. The second derivative tells us how dP/dt is changing at any instance in time, and it has units of rabbits per month squared.

Ok, so if a function f which takes in units u and outputs units v we have the following chart for any order of derivative:

order of derivative	Output Units
0	v
1	v/u
2	v/u^2
3	v/u^2
$\vdots$	$\vdots$
n	v/u^n

so we can always interpret higher derivatives as telling us how a derivative one step lower is changing. But what is going on mathematically, and how can it help us with optimization?

Recall, that if f was a function, and a was real number, then the tangent line at a is given by:

$$l(x) = \frac{df}{dx}(a)(x - a) + f(a)$$

Now note, that at a we have:

$$l(a) = f(a) \qquad \text{and} \qquad \frac{dl}{dx}(a) = \frac{df}{dx}(a)$$

The reason that l is such a good approximation for f at $x = a$ is precisely because at a $l(a)$ is $f(a)$, and $dl/dx(a)$ is $df/dx(a)$ at a. In other words, l is the unique line which at a best mimics the behavior of f as it is equal to f at a and changing in the same way as f at a. However, this approximation is only so good, indeed d^2l/dx^2 is identically zero, while $d^2f/dx^2(a)$ definitely does not need to be zero.

How can we upgrade our approximation of f with a simple function? Well we want to stick with polynomials as they have the simplest non trivial derivatives, and if we want a better approximation of f we will need the second derivative of f to agree with the second derivative of our approximation; otherwise we are doing no better than the linear case. Putting all of this together, it follows that we want a quadratic function $p(x)$ so that at $x = a$ we have:

$$p(a) = f(a) \qquad \frac{dp}{dx}(a) = \frac{df}{dx}(a) \qquad \frac{d^2p}{dx^2}(a) = \frac{d^2f}{dx^2}(a)$$

It follows that p should be of the form:

$$p(x) = \alpha(x-a)^2 + \frac{df}{dx}(a)(x-a) + f(a)$$

as at a we have that:

$$p(a) = \alpha(a-a)^2 + \frac{df}{dx}(a)(a-a) + f(a) = 0 + 0 + f(a) = f(a)$$

and:

$$\frac{dp}{dx} = 2\alpha(x-a) + \frac{df}{dx}(a) \Rightarrow \frac{dp}{dx}(a) = \frac{df}{dx}(a)$$

But what should α be? Well taking a second derivative, we have that:

$$\frac{d^2p}{dx^2} = 2\alpha$$

which we want to be equal to $d^2f/dx^2(a)$, hence we have that:

$$\alpha = \frac{1}{2}\frac{d^2f}{dx^2}$$

so our $p(x)$ should be given by:

$$p(x) = \frac{1}{2}\frac{d^2f}{dx^2}(a)(x-a)^2 + \frac{df}{dx}(x-a) + f(a)$$

We call $p(x)$ the *tangent parabola to f at a*.

Example 6.6. Let $f(x) = e^x$, we want to find the tangent parabola to e^x at $x = 0$. Well:

$$\frac{df}{dx} = e^x \qquad \text{and} \qquad \frac{d^2f}{dx^2} = e^x$$

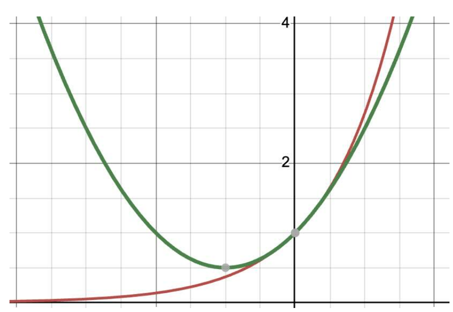
hence at $x = 0$

$$\frac{df}{dx}(0) = 1 \qquad \text{and} \qquad \frac{d^2f}{dx^2}(0) = 1$$

so the tangent parabola is given by:

$$p(x) = \frac{1}{2}x^2 + x + 1$$

We have the following picture:



Example 6.7. Let $f(x) = \sin x$, we want to find the tangent parabola to $\sin x$ at $x = \pi/4$. Well:

$$\frac{df}{dx} = \cos x \qquad \text{and} \qquad \frac{d^2 f}{dx^2} = -\sin x$$

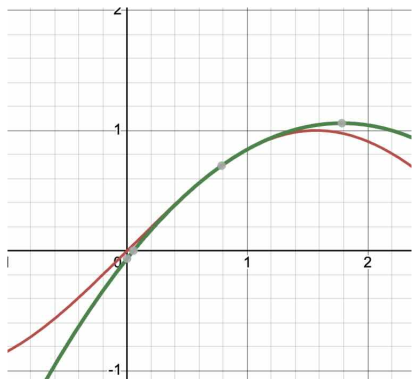
hence at $x = 0$:

$$\frac{df}{dx}(0) = \frac{\sqrt{2}}{2} \qquad \text{and} \qquad \frac{d^2 f}{dx^2}(0) = -\frac{\sqrt{2}}{2}$$

so:

$$p(x) = -\frac{\sqrt{2}}{4}(x - \pi/4)^2 + \frac{\sqrt{2}}{2}(x - \pi/4) + \frac{\sqrt{2}}{2}$$

We have the following picture:



6.3 Concavity and the Second Derivative Test

We do todays warm up:

Example 6.8. The formula for the tangent parabola is given by:

$$p(x) = \frac{1}{2}\frac{d^2 f}{dx^2}(a)(x - a)^2 + \frac{df}{dx}(x - a) + f(a)$$

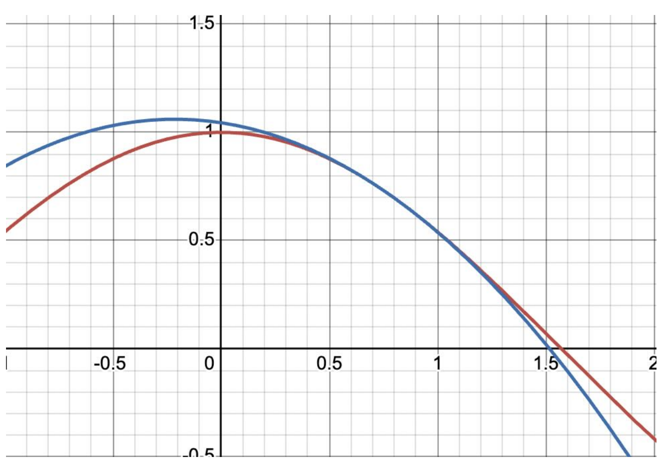
For $\cos x$ we have that:

$$\frac{d \cos x}{dx} = -\sin x \qquad \text{and} \qquad \frac{d^2 \cos x}{dx^2} = -\cos x$$

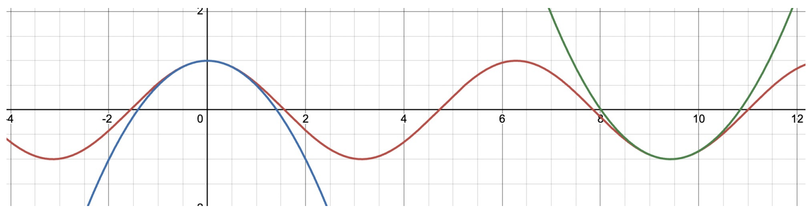
hence at $x = \pi/4$ we have that:

$$p(x) = -\frac{\sqrt{2}}{4}(x - \pi/4)^2 - \frac{\sqrt{2}}{2}(x - \pi/4) + \frac{\sqrt{2}}{2}$$

Here is a picture of the situation:



Note that the sign of the first derivative told us whether or not a function one was increasing or decreasing, but what does the sign of the second derivative tell us? Well, the tangent parabola at a point incorporates the curvature of the function into our approximation. In particular, whichever direction the parabola is facing, either upwards or downwards, matches whichever direction the function is curving in:



However, the direction in which a parabola points is entirely determined by the sign of it's leading order coefficient, i.e. $x^2 + 3x + 12$ points up, but $-x^2 + 3x + 12$ points down. It follows that the sign of the second derivative tells us two pieces of information: firstly, since the second derivative is the derivative of the derivative, the sign of the second derivative tells us whether the derivative itself is increasing or decreasing. Secondly, the sign of the second derivative tells whether our function is curving up (positive), or curving down (negative). This is such an important concept we give it a name:

Definition 6.1. Let f a function, then f is **concave up** at a if $d^2f/dx^2(a) > 0$, and f is **concave down** at a if $d^2f/dx^2(a) < 0$. We say that a is a **inflection point** if d^2f/dx^2 switches sign at a (in particular this implies that d^2f/dx^2 is either equal to zero or does not exist).

One should always remember that concave up means the function is facing up, and concave down means the function is facing down. We check some simple examples of this:

Example 6.9. The function e^x is always concave up, as is e^{-x} even though e^{-x} is decreasing.

Example 6.10. Any linear function is neither concave up nor concave down. A quadratic function is constantly concave up or concave down depending on the sign of the leading coefficient. More concretely, we have that if:

$$f(x) = 3x^2 + 2x - 12$$

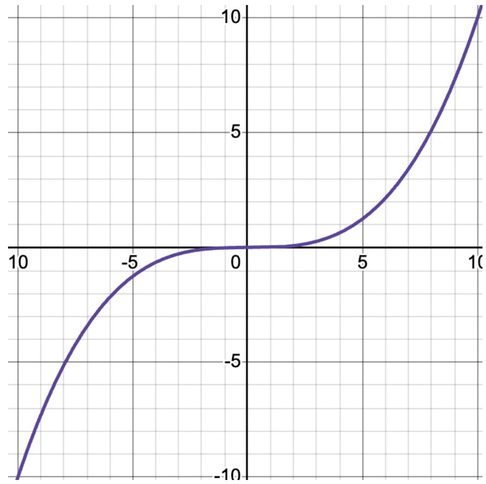
then the second derivative is the constant function $d^2f/dx^2 = 6$. If instead:

$$f(x) = -3x^2 + 2x - 12$$

then the second derivative is the constant function $d^2f/dx^2 = -6$. Just as linear functions are those which are constantly increasing or decreasing, quadratics are those which are constantly concave up or concave down.

Finding inflection points is basically the same procedure as finding critical point, and checking to see if they are local minima and local maxima.

Example 6.11. We consider the function x^3. Then the second derivative is $6x$, which is zero when $x = 0$. When $x < 0$ we have that x^3 is concave down, and when $x > 0$ we have that x^3 is concave up, so 0 is an inflection point. We have the following picture:



Figuring the general shape of the graph of a function is not all that second derivative is good for, indeed we can use it for optimization purposes as well. If a is a critical point of f, and $d^2f/dx^2(a) > 0$, then at a, $f(x)$ looks like the bottom of a parabola facing upwards, i.e. a looks like a local minimum. Similarly, if a is a critical point of f and $d^2f/dx^2(a) < 0$, then at a, $f(x)$ looks like the bottom of a parabola facing upwards, i.e. a looks like a local maximum. One can also phrase this as follows: if $d^2f/dx^2(a) < 0$ (or $d^2f/dx^2(a) > 0$), then df/dx is decreasing (or increasing) at a, so if a is a critical point then df/dx must be going from positive to negative (or negative to positive), and so a is a local maximum (or local minimum).

This is called the second derivative test, and we enshrine it with a theorem:

Theorem 6.1. *Let f be a function, and a a critical point of f. Then if $d^2f/dx^2 > 0$ then a is a local minimum, and if $d^2f/dx^2 < 0$ then a is a local maximum.*

Example 6.12. Let:

$$f(x) = x^3 - 6x$$

This function has first derivative:

$$\frac{df}{dx} = 3x^2 - 6$$

and thus has critical points $x = \pm\sqrt{2}$. This function has second derivative:

$$\frac{d^2f}{dx^2} = 6x$$

and so $-\sqrt{2}$ is a local maximum, while $\sqrt{2}$ is a local minimum.

Note that the second derivative, while computationally less intensive than using the first derivative, does not always tell us whether a critical point is a minimum or a maximum. Indeed, consider the following example:

Example 6.13. Let:

$$f(x) = x^4$$

then common sense tells us we have a minimum at 0, and indeed, the first derivative is zero at 0 as:

$$\frac{df}{dx} = 4x^3$$

and the derivative switches from negative to positive at 0, hence 0 is a local minimum. However, the second derivative cannot tell us whether 0 is a local minimum or maximum as:

$$\frac{d^2 f}{dx^2} = 12x^2$$

hence $d^2 f/dx^2$ is 0 at 0, which tells us nothing.