Calculus I: Lecture Notes

Quinn La Fond

April 6, 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 $\square$. 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:

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

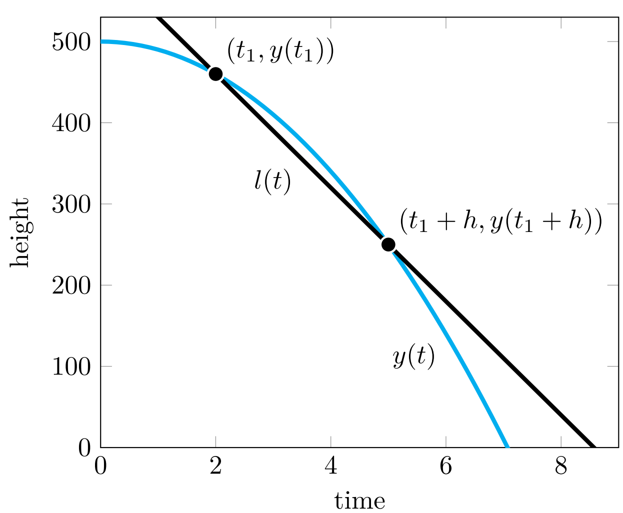
Now, v_{avg} is the slope of the line which goes through the points:

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

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

[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.

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

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:

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

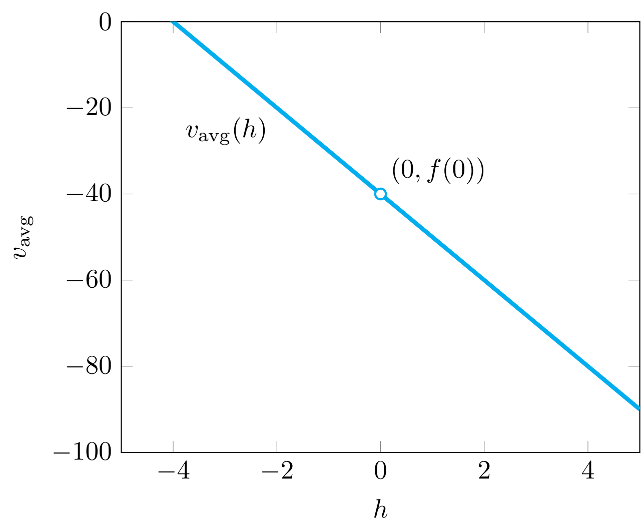
which doesn't make mathematical or physical sense! So, how can we rectify the situation? The next key insight is that if we allow h to vary instead of being fixed, then as h gets closer and closer to zero, since the interval becomes smaller and smaller, we get average velocities which are closer and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

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

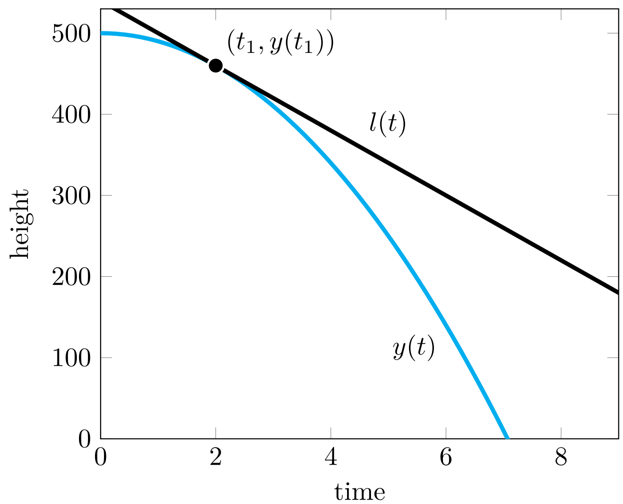
This is a *rational function*, and as such is not defined at $h = 0$, however, for all $h \neq 0$, we have that:

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

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

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

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.

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

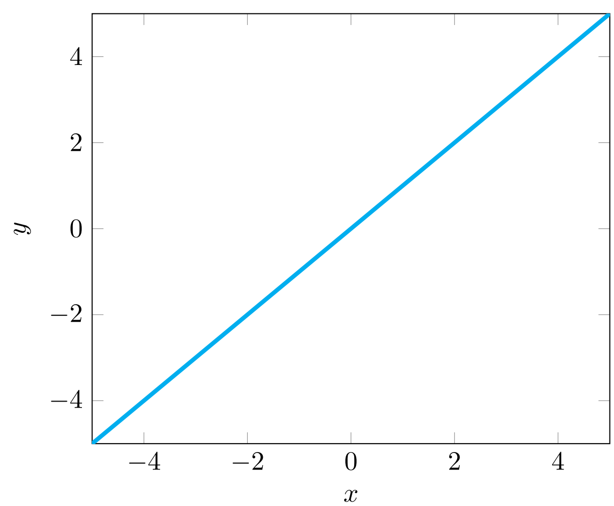
Limits essentially come in two flavors, we have limits as 'x goes to positive or negative infinity', and limits as 'x approaches a', where a is some finite real number. The former is a way of characterizing the long term behavior of a function $f(x)$, and the later analyzes the behavior of f near a point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



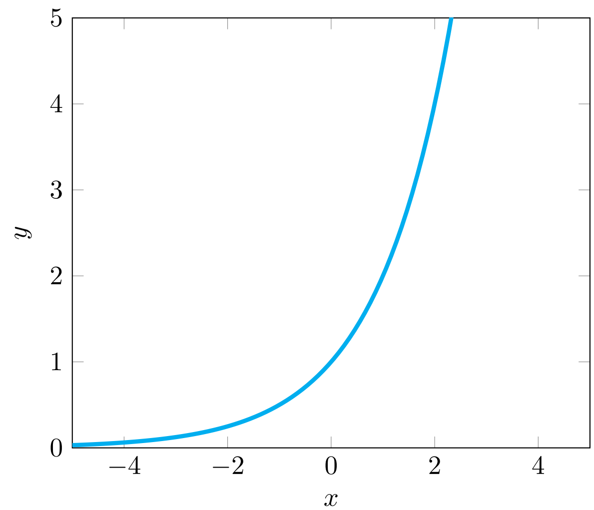
Now as we plug larger and larger and numbers into $f(x)$, we just return that same number since $f(x) = x$. In particular as x gets bigger $f(x)$ bigger, so the long term behavior of $f(x)$ is to get bigger and bigger. When this occurs, we write:

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

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:*

$$p(x)=a_nx^n+\cdots+a_1x+a_0 \qquad \textit{and} \qquad q(x)=b_mx^m+\cdots+b_1x+b_0$$

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

$$\lim_{x\to\infty}\frac{p(x)}{q(x)}=\lim_{x\to\infty}\frac{a_nx^n}{b_mx^m} \qquad \textit{and} \qquad \lim_{x\to-\infty}\frac{p(x)}{q(x)}=\lim_{x\to-\infty}\frac{a_nx^n}{b_mx^m}$$

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:

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

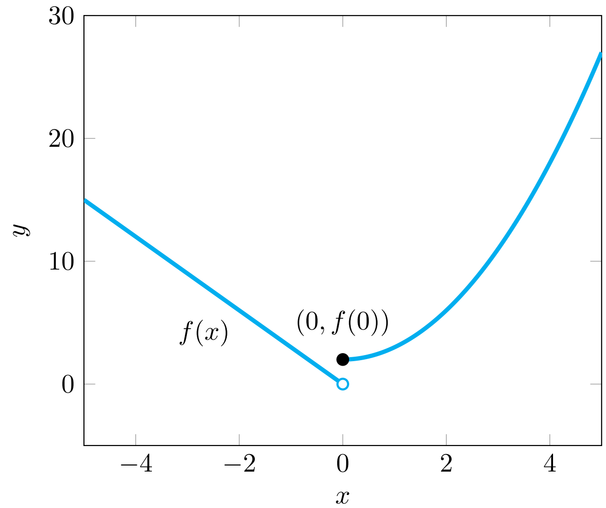
The next situation is that:

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

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.

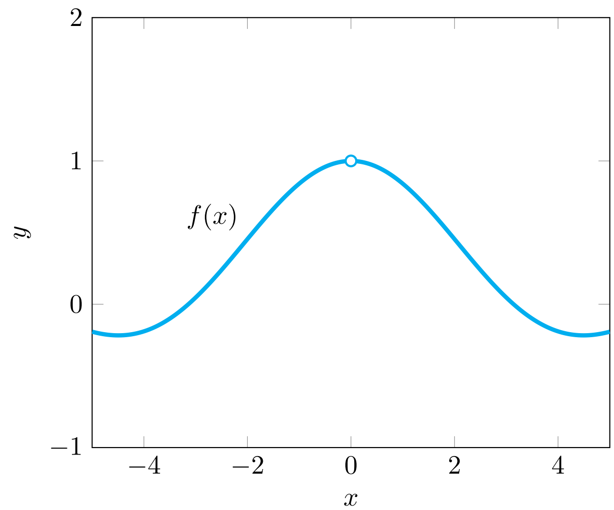
b) If $\lim_{x \to a^-} = \lim_{x \to a^+} = L$, ∞ or $-\infty$, then $\lim_{x \to a} f(x) = L$, ∞, or $-\infty$ respectively.

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:

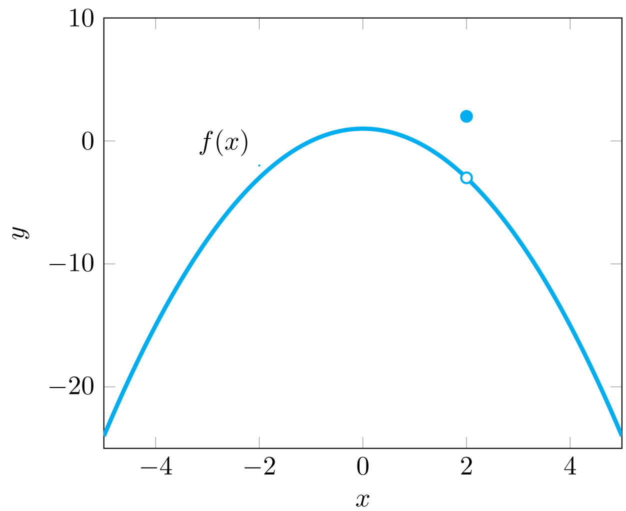
$$f(x) = \begin{cases} -x^2+1 & \text{if } x \neq 2 \\ 2 & \text{if } x = 2 \end{cases}$$

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.

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

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 $f(x)$.

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

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

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

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

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

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:

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

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 $\lim_{x \to a} \sin x = \sin a$. 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:

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

It follows that for all x:

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

[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.

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]

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

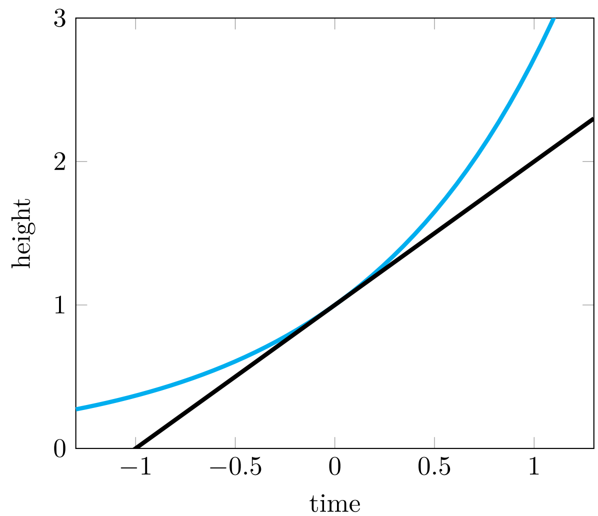
Example 1.14. The following function will be of great importance to us in the future:

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

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]

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

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:

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

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

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 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:

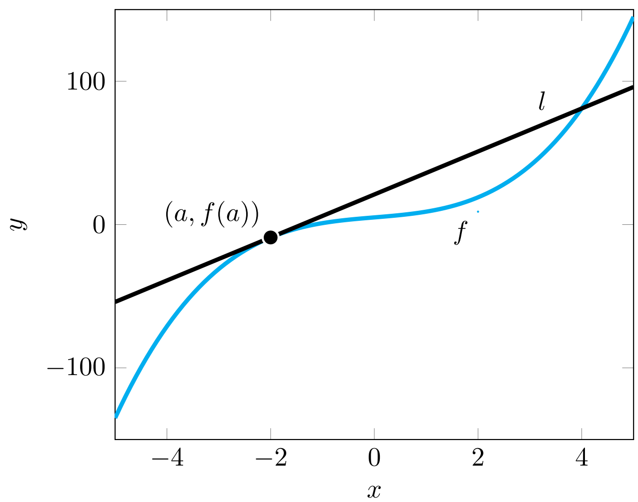
$$\lim_{h\to 0^-}\frac{f(h)-f(0)}{h}=\lim_{h\to 0^-}\frac{-h}{h}=-1$$

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 $x = 0$.

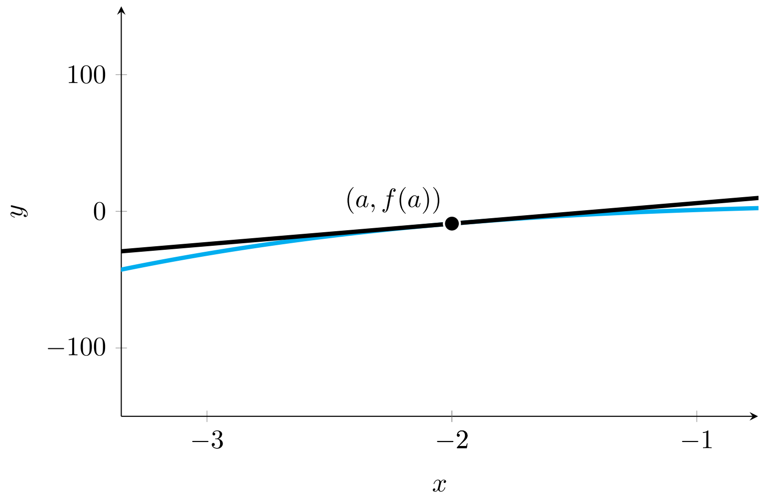
Before calculating more examples of functions, we wish to provide a geometric interpretation of the derivative. Our original motivation came from physics, where we thought of the derivative of a distance function as giving us a velocity function; but in full generality what is the derivative actually 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:

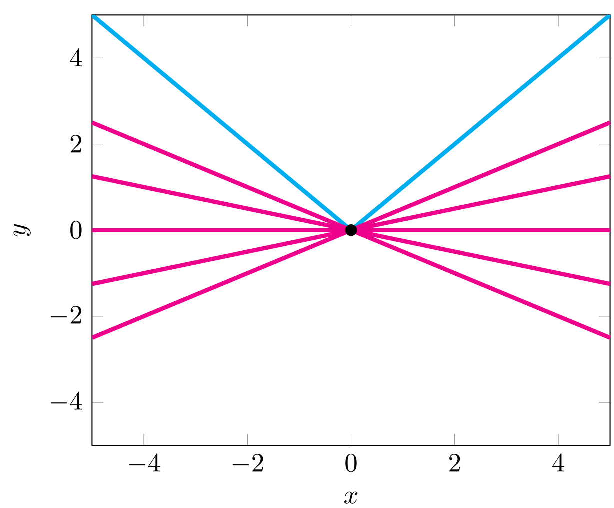
$$l(x) = f'(a) \cdot (x - a) + f(a)$$

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

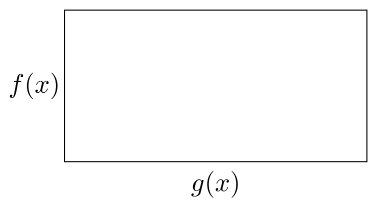
In the previous section, we were able to easily find that the derivative of a sum functions is the sum of their derivatives, and that the derivative of a real number times a function was that real number multiplied with that function. Symbolically this is stated as:

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

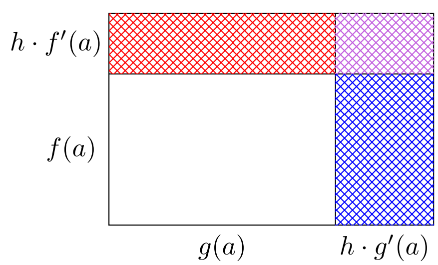
In this section we explore what happens when we try to take derivatives of products and quotients. One may initially expect that:

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

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:

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

[12] See next section

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:

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

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:

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

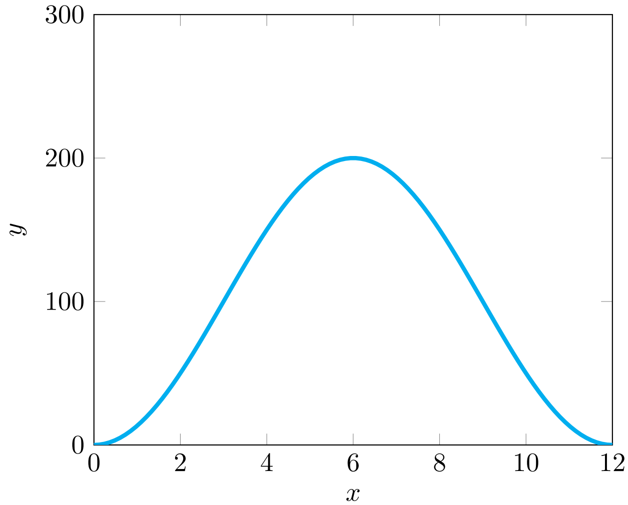
Abstractly thinking of one function as a function of another function is hard, so let us come up with a reasonable example.

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

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:



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:

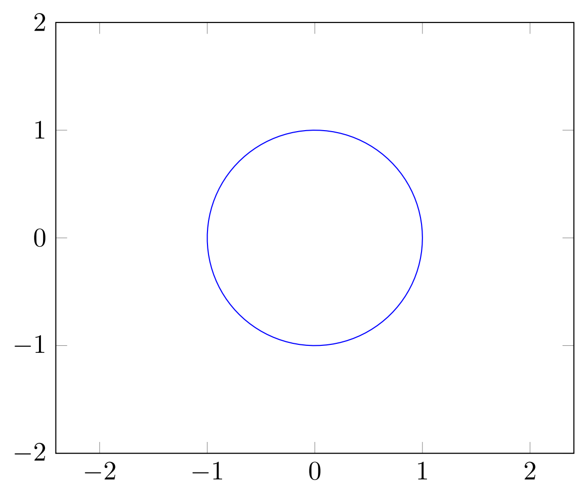
$$\begin{aligned}\frac{dF}{dt} &= \frac{dF}{dr} \cdot \frac{dr}{dt} \\ &= -\frac{2k}{r^3} \cdot \frac{1}{t+1} \\ &= -\frac{2k}{(\ln(t+1))^3} \cdot \frac{1}{t+1}\end{aligned}$$

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:

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

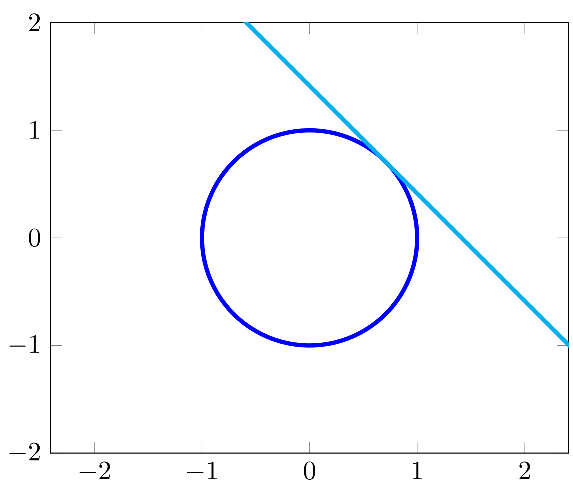
Now we can plug in $x = \sqrt{2}/2$ and $y = \sqrt{2}/2$ to solve for dy/dx:

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

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