# A Calculus Companion

Marlin Figgins

August 17, 2018

# Introduction

These notes are intended to be used as an introduction to calculus. Ideally this is meant to accompany a calculus course like that of MATH 13100-13200-13300 or MATH 15100-15200-15300 at the University of Chicago but can also be used for independent study. During my time as an undergrad at the University of Chicago, I was a TA for 6 quarters of the MATH 130s sequence. What I learned from my students was that many of them thought they were bad at mathematics. Increasingly, I began to see a divorce between theory, example, and application in the way the course was taught. My goal in these notes is to present calculus in a way that emphasizes both its reason for existing and its usefulness. I hope that by the end of this, the reader will be able to understand the language of calculus and its appreciate its usefulness. Most importantly, I wish to enable my reader to apply the ideas of this text themselves to problems they may encounter in the future.

Therefore, I will try my best to balance the theory behind calculus with my use of examples. I admit that my main purpose is to motivate calculus as a tool and develop the theory in accessible language. Most importantly, I hope provide the intuition necessary to understand calculus and its various concepts, so there may not be a ton of examples outside of those necessary to emphasize the thought process behind solving problems with calculus. In the future, I will provide exercises at the end of each section!

# Table of Contents

# 1.    What is calculus?

The word *calculus* comes from Latin, literally meaning 'small pebble' which frankly isn't too far off from the actual concerns of calculus. Simply put, calculus asks two main questions:

**Question 1.1** (The Two Main Questions of Calculus)**.**

     *1. What does a small change in one thing tell you about change in something else?*

     *2. How do these small changes add up to give a sum?*

We will approach the first of these questions using *differentiation* and the second with another concept called *integration.* As we will see, these two questions are intimately related, and we'll be able to find a very useful relationship between differentiation and integration.

Before we begin trying to unpack and answer these two questions, I'm sure you must be forming two of your own?

**Question 1.2.** *Why am I learning this? Why is any of this useful?*

Differentiation and integration have numerous applications in statistics, physics, chemistry, biology, economics, and generally are useful for understanding how things change in the world around us. In fact, some of the things in the world around you (like your brain, car, and phone) are already using calculus and its principles which we will explore through caes studies.

Calculus is used to calculate energy and motion in physics, reaction rates and radioactive decay in chemistry, birth and death rates in biology, and costs and revenues in economics. In many cases, calculus is the natural language for describing a problem. Let's start with one such example where the central ideas of calculus naturally appear.

**Example 1.3.** Suppose two people $M$ and $J$ are having a race from a point $a$ to another point $b$. When watching this sort of thing play out, we develop a vague idea or prediction of the winner is going to be. In order to come to a conclusion, we unconsciously ask ourselves questions like "Where are $M$ and $J$?" and "How fast are they running?". To answer the first, we simply recall where they were at a particular time $t^*$. This method doesn't work as well when it comes to the second, how can we remember their speed right at time $t^*$? Let's take a step back and look at where they were at different points in the race.

If we look at Example 1.3, we see an example of a race like this. In this case, $J$ wins the race in 5 seconds and $M$ loses it, so intuitively, $J$ must have moved faster than $M$. We can attempt to capture this using the average speed formula

$$\text{Average Speed} = \frac{\text{Distance Traveled}}{\text{Time}}. \tag{1.4}$$

If we call the speed of $M$ $v_M$ and the speed of $v_J$, then $M$ and $J$ have speeds

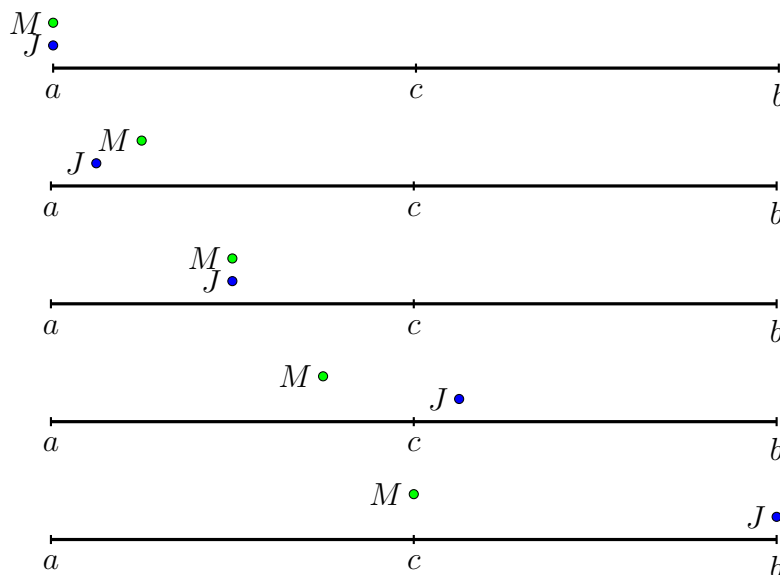$$v_M = \frac{c - a}{5} \text{ and } v_J = \frac{b - a}{5}. \tag{1.5}$$

Figure 1: Visualizing the race

This confirms our intuition that $J$ was faster at the end of the day, but something is missing here. At the start of the race, $J$ is behind $M$ and moving slower, so $J$ must have began to speed up. Therefore, it's incorrect to say that $J$ is moving at speed $v_J$ at every point in time. Let's try a different approach. First, we visualize the race by graphing both runners' position at every point in race.

From this, we see that the positions of $M$ and $J$ are functions of time. If we're given a point in the race $t$, we can find their positions $x_M(t)$ and $x_J(t)$ using functions $x_M$ and $x_J$. This allows us to rephrase the question. If we want to figure out how fast $M$ and $J$ moving at $t^*$, we need to see how their position functions change around that time. In short, we want the rate of change of $x_M$ and $x_J$ around $t^*$. In algebra 1 or 2, you learn to calculate rate of change as the slope $m$ of a line $y = mt + b$.

$$m = \frac{\text{Change in } y}{\text{Change in } t} = \frac{\Delta y}{\Delta x}.$$

We'll write these changes as $\Delta y$ and $\Delta t$. If we zoom in on our functions $x_J$ and $x_M$ far enough, we see that they look straight lines. This allows us to look at small differences in $t$, $x_M$, and $x_J$. Let's say $x_M(2.1) = 4.2$, $x_M(1.9) = 3.8$, $x_M(2.1) = 4.41$, and $x_M(1.9) = 3.61$. If we try to approximate the rate of change this way, then the velocity of $M$ and $J$ at $t^* = 2$ are approximately given by:

$$v_M^* \approx \frac{\Delta x_M}{\Delta t} = \frac{4.2 - 3.8}{2.1 - 1.9} = \frac{0.4}{0.2} = 2 \tag{1.6}$$

$$v_J^* \approx \frac{\Delta x_J}{\Delta t} = \frac{4.41 - 3.61}{2.1 - 1.9} = \frac{0.8}{0.2} = 4. \tag{1.7}$$
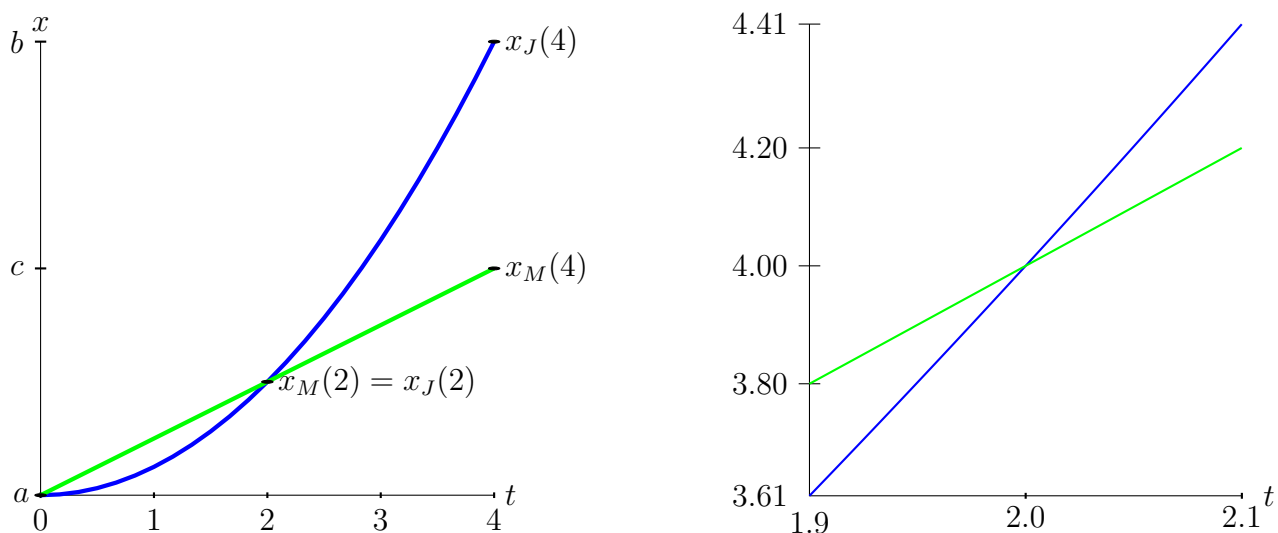
Figure 2: Graph of time $t$ v. position $x$ at different scales.

We now know that if we zoom in very close and look at small differences, then we're able to approximate how something is changing at a particular moment. This is the same idea behind how a car's speedometer calculates the speed of the car. Calculus comes in when we try to make this method more generally applicable.

# 2.    Limits and Continuity

In the previous example, we saw that zooming in and only looking at the average speed for small time intervals can help us figure out the speed at a particular moment in time, but we still don't know exactly how close we have to zoom to guarantee the right answer. This leads us to the concept of limits.

Conceptually a limit asks, *what happens to our function when you get very close to a certain point?*. Looking at our previous example, we zoomed in on $t^*$ to find the limit of $\frac{\Delta x}{\Delta t}$ as $\Delta t$ got small or *approaches 0*. This is idea behind a limit. Since the position of the runners changed in time, we were able to 'zoom in' on a particular time to find how fast they were moving at that instant.

## 2.1    Preliminaries & Limit Definitions

Before moving for let's look at this more abstractly and develop some notation for dealing with these questions. From the discussion above, we can see that we're interested in functions of some kind of *variable* like time, position, height, energy, population, dollars or whatever quantity you might think of. That is, if we have a number $x$ representing a quantity like this, we want to assign it to another number $f(x)$. We'll call the rule $f$ which takes $x \mapsto f(x)$ a *function*. Usually, we'll just write $x \in \mathbb{R}$ to say that $x$ is a *real number* like the ones we encounter in everyday life. $1,2,3$, $\frac{2}{3}$, $1.21$, and $3.14$ for example. With this in mind, since the function $f$ takes real numbers to real numbers, we can write it as

$$f \colon \mathbb{R} \to \mathbb{R}.$$

Alternatively, if $f$ isn't defined for all numbers $\mathbb{R}$, but only some subset of numbers $A$, we can write that $f \colon A \to \mathbb{R}$. This means that $f$ takes $x$ in $A$ to $f(x) \in \mathbb{R}$. These sets $A$ are usually 'nice' sets like *open interval* $(a,b)$ i.e. the set of real numbers that are between $a$ and $b$ or

$$(a,b) = \{x \in \mathbb{R} \mid a < x < b\}$$

or the closed interval which includes the endpoints

$$[a,b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}.$$

**Remark 2.1.** *The brackets $\{\cdot\}$ used above are used to denote a* set *and we often use set notion to denote a 'container' of objects as a way of simplify our notion. For example, we'll write $\{y \mid C\}$ to denote the collection of objects $y$ that satisfy a condition $C$. We'll say that $x \in \{y \mid C\}$ to mean that $x$ is an element of our set i.e. that $x$ satisfies the condition $C$.*

Let's we reframe this limit idea in this fancy new notation. If we have a function $f \colon A \to \mathbb{R}$ and want to find the limit as $x$ approaches $a$ to exist, then we want $f(x)$ to approach a single number as we get close to our point $a$. In order to get at what this means mathematically,

Figure 3: $|x - 4| < 7$ on the number line.

we need something called the *absolute value* or *norm*, which is what mathematicians use to measure magnitudes and distances. As you will see, the absolute value follows most of the same rules you would expect of distances in real life.

**Definition 2.2** (Absolute Value). The absolute value of a number $a$ is defined by:

$$|a| = \begin{cases} a & \text{if } a \geq 0 \\ -a & \text{if } a < 0. \end{cases} \tag{2.3}$$

We can think as the absolute value as function that takes in a number and returns the positive part. For example, $|-3| = 3$. In this way, it is similar to distance. In real life, if you move 5 feet to the left or to the right, you still moved 5 feet; the only difference is *direction*. Therefore, the absolute value is only concerned with the *magnitude* or *total distance* traveled by saying that moving 5 feet in any direction is the same in terms of distance. Similarly, we say that the distance between two points $x$ and $a$ is given by $|x - a|$. We can use inequalities to illustrate this. What if I ask you to solve the inequality

$$|x - 4| < 7.$$

We can solve this by noticing that if $|x - 4| < 7$, then both $x - 4 < 7$ and $4 - x < 7$, so $x - 4 > -7$. Doing some work with inequalities, we see that

$$-7 < x - 4 < 7 \implies -3 < x < 11. \tag{2.4}$$

We can visualize this on a number line.

The absolute value has several interesting properties that are consistent with our interpretation of it as distance.

**Proposition 2.5** (Properties of Absolute Value). *For any numbers $x$ and $y$*

1. *$|x| = 0$ if and only if $x = 0$,*
2. *$|x| = |-x|$,*
3. *$|xy| = |x| \, |xy|$,*
4. *$|x^2| = |x|^2 = x^2$,*
5. *$|x + y| \leq |x| + |y|$.*

The last property is called the *triangle inequality* and will prove to be very useful in establishing some of the properties of limits.

With an idea of distance and closeness at hand, we can revisit the question of *what happens to our function when you get very close to a certain point?*. Let's say we have a function $f \colon A \to \mathbb{R}$ and a point $a$. If we want to find the *limit of $f$ as $x$ approaches $a$*, we want to find a number $L \in \mathbb{R}$ such that $f(x)$ *approaches $L$ when $x$ approaches $a$*. If we can find a number like this, we write

$$\lim_{x \to a} f(x) = L. \tag{2.6}$$

**Definition 2.7.** Formally, we define this by

$$\lim_{x \to a} f(x) = L \tag{2.8}$$

if for every $\epsilon > 0$, there exists $\delta > 0$ such that

$$|x - a| < \delta \implies |f(x) - L| < \epsilon.$$

**Remark 2.9.** *In words, we would say 'the limit of $f$ as $x$ approaches $a$ is equal to $L$'.*

This is an example of a '$\epsilon$-$\delta$' statement which can be very difficult to unpack when you first see.To help give the notation above meaning, we attempt to visualize this $\epsilon$-$\delta$ definition of a limit in Section 2.1 below.

Intuitively, this is saying that if $\lim\limits_{x \to a} f(x) = L$, then our function $f$ can get close as we want to some value $L$ if we zoom in close enough to $a$ i.e. if $f(x) \approx L$ whenever $x \in (a - \delta, a + \delta)$ for some small number $\delta$.

Be fair warned, not all functions have proper limits everywhere! Take

$$g(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0, \end{cases}$$

for example! If we graph this function, we'll see that it clearly has no limit at $0$. Before moving on to some of the interesting properties of limit, let's show an example of how to use the $\epsilon - \delta$ definition to prove a limit exists.

**Example 2.10.** Calculate $\lim\limits_{x \to 2}(3x - 1)$ and prove your result.

Looking at the graph of $f(x) = 3x - 1$, you can see that $\lim\limits_{x \to 2} 3x - 1 = 5$, but in order to *prove* this, we must use the $\epsilon$-$\delta$ definition.

Suppose we have some small number $\epsilon > 0$. If the limit exists, then we know that if we zoom in close enough to $a = 2$ on the $x$-axis that our function $f(x) = 3x - 1$ will be $\epsilon$ close to the limit $L = 5$. The point of $\delta$ is to tell us how close we need to be to $a = 2$
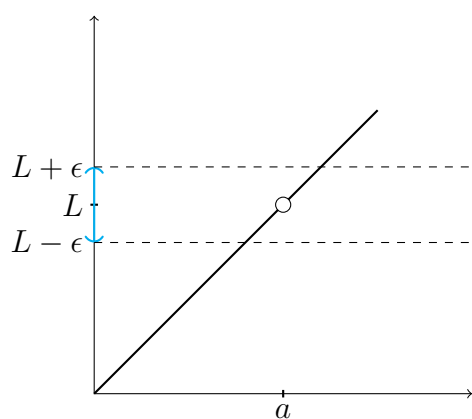
(a) For every $\epsilon > 0$,
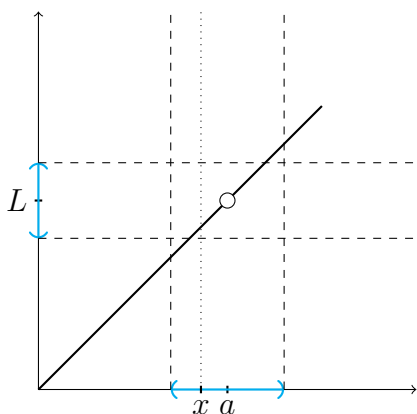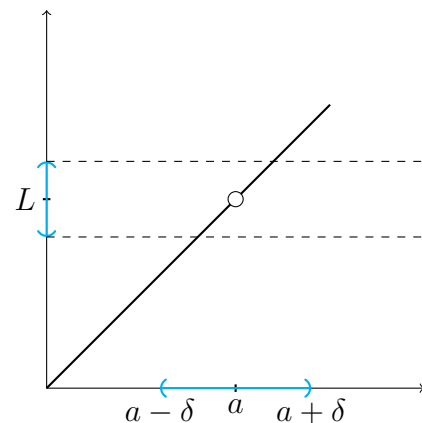
(b) there exists $\delta > 0$ such that

(c) if $|x - a| < \delta$,

(d) then $|f(x) - L| < \epsilon$.

Figure 4: Visualizing the $\epsilon$-$\delta$ definition.

Figure 5: Visualizing the argument above.

so that $f(x) = 3x - 1$ is $\epsilon$-close to 5. Let's show this mathematically in terms of Definition 2.7.
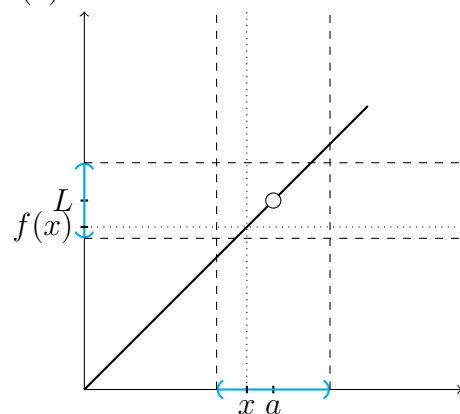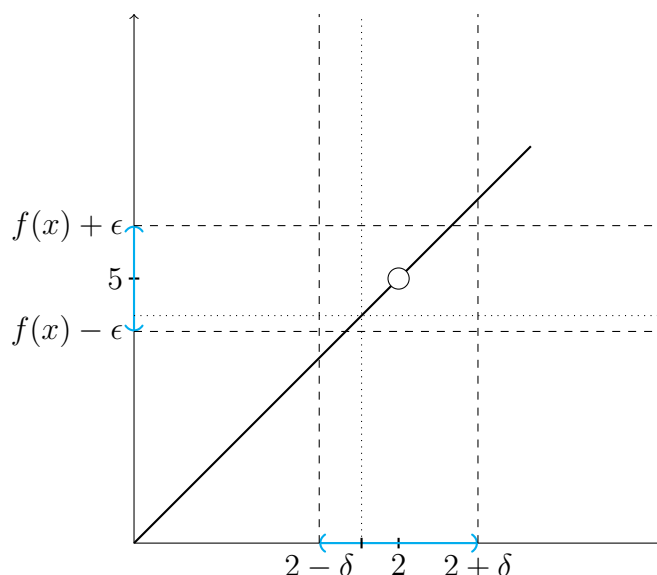
We see that if $f(x)$ is $\epsilon$-close to 5, we can write this mathematically as

$$|f(x) - 5| < \epsilon.$$

If we substitute in the definition of $f(x)$, we see

$$|(3x - 1) - 5| = |3x - 6| < \epsilon,$$

but what we want to find is a number $\delta$, so that if $x$ is $\delta$-close then $f(x)$ is $\epsilon$-close to 5! We can write $x$ is $\delta$-close to 2 as

$$|x - 2| < \delta.$$

If you compare the two inequalities above, you see that you can make the lefthand sides equal by multiplying by 3 which shows

$$3\,|x - 2| = |3x - 6| = |3x - 1) - 5| < 3 \cdot \delta.$$

From this, we see that if $\delta = \frac{\epsilon}{3}$, then

$$|3x - 6| < 3 \cdot \delta = 3 \cdot \frac{\epsilon}{3} = \epsilon,$$

which gives us the inequality we want to see! With this information, we can write the proof as follows.

*Proof.* Let $\epsilon > 0$ and $\delta = \frac{\epsilon}{3}$. If

$$|x - 2| < \delta,$$

then after multiplying by 3, we see

$$|3x - 6| < \epsilon.$$

Therefore,

$$|3x - 6| = |(3x - 1) - 5| = |f(x) - 5| < \epsilon,$$

which proves that the limit is equal to 5. □

The idea of limits is conceptually robust and natural, but when it comes to most of applications we're interested in, we'll encounter more complicated functions which make this approach either extremely difficult or just tedious. Because of this, we'll develop some methods of computing complicated limits using simpler functions which we already know the limit of. For example, if we have two functions $f$ and $g$ which have limits $L$ and $M$, it's reasonable to assume that $f + g$ converges to $L + M$ (why?). Is this also true for $f \cdot g$ or $10 \cdot f$? We'll provide an answer to these questions with the following with the following theorems.

## 2.2　Limit Theorems

**Theorem 2.11** (Properties of Limits)**.** *If the limits of two functions $f(x)$ and $g(x)$ exist at $a$. That is, if $\lim_{x \to a} f(x) = L$ and $\lim_{x \to a} g(x) = M$ for some numbers $L$ and $M$, then*

*1.* $\lim_{x \to a} [f(x) + g(x)] = L + M,$

*2.* $\lim_{x \to a} [cf(x)] = cL$ *if $c$ is any number,*

*3.* $\lim_{x \to a} [f(x) \cdot g(x)] = L \cdot M,$

*4. if $M \neq 0$, $\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{L}{M}$.*

As you can see, these rules make calculating limits like

$$\lim_{x \to 3} \frac{x^2 - 1}{x^8 - 3x} \tag{2.12}$$

much easier. In this case, it's as simple as plugging in 3 into our equation, but it won't always be this easy. We'll formalize this idea of functions are 'nice' and 'not-so-nice' to evaluate with limits in the next sections as we develop the ideas of *continuity* and *differentiblity*.

Conceptually, we say that a function is continuous if we 'can draw its graph without lifting our pencil', but mathematically this means that if we move a tiny little on the $x$-axis then we also move a tiny bit up or down on the graph of our function. Let's try and write this idea in terms of limits.

**Definition 2.13.** We say that a function $f(x)$ is *continuous at a* if

$$\lim_{x \to a} f(x) = f(a). \tag{2.14}$$

If $f(x)$ is continuous at every point it is defined, then we just say that $f(x)$ is continuous.

As you might expect, not all functions are continuous, take $\frac{1}{x}$ for example. This fails to be continuous at 0.

The limit properties described earlier are useful for determining continuity.

**Theorem 2.15.** *If $f$ and $g$ are continuous at $a$, then*

*1. $f + g$, $f \cdot g$ are continuous at $a$,*

*2. $cf$ is continuous at $a$ for any real number $c$,*

*3. $\frac{f}{g}$ is continuous at $a$ whenever $g(a) \neq 0$.*

We also have another property for the composition of continuous functions.

**Theorem 2.16.** *If $f$ is continuous at $a$ and $g$ is continuous at $f(a)$, then $g \circ f$ is continuous at $a$.*

There are other properties of limits and continuity such as the squeeze and intermediate value theorems! We'll begin with the squeeze theorem.

**Theorem 2.17** (Squeeze Theorem). *Suppose $f(x)$ is defined near $c$, and there are functions $U(x)$ and $L(x)$ which bound $f(x)$ near $c$ i.e.*

$$L(x) \leq f(x) \leq U(x) \text{ for every } x \text{ near } c.$$

*If $\lim_{x \to c} U(x) = \lim_{x \to c} L(x)$, then*

$$\lim_{x \to c} L(x) = \lim_{x \to c} f(x) = \lim_{x \to c} U(x).$$

This is the statement that if $f(x)$ is sandwiched between two functions $L(x)$ and $U(x)$ that approach each other, then $f(x)$ must approach their common value. We can see an application of this theorem in the following example.

**Example 2.18.** What is the limit of $f(x) = x \sin\left(\frac{1}{x^2}\right)$ as $x$ approaches 0?

Graphing the function ( Section 2.2), we see that as we zoom into the origin, $f(x)$ appears to approach zero. We can see this analytically by considering the absolute value of $f(x)$ as follows
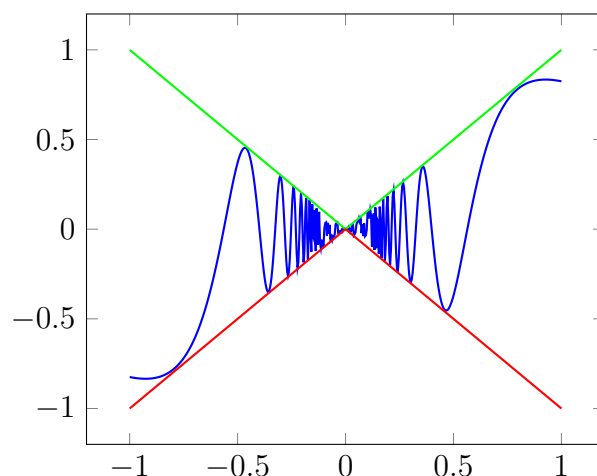
Figure 6: Graphs of $f(x) = x \sin\left(\frac{1}{x^2}\right)$, $L(x) = -|x|$, $U(x) = |x|$.

$$|f(x)| = \left| x \sin\left(\frac{1}{x^2}\right) \right| \tag{2.19}$$

$$= |x| \cdot \left| \sin\left(\frac{1}{x^2}\right) \right|. \tag{2.20}$$

The part of interest here is the $\sin\left(\frac{1}{x^2}\right)$. One might remember from pre-calculus that $\sin(\theta)$ is bounded between 1 and $-1$ for every number $\theta$. Therefore, we see that

$$|f(x)| \leq |x| \text{ for any } x \text{ near } 0.$$

Breaking down the absolute value around $f(x)$, then gives us

$$-|x| \leq f(x) \leq |x| \text{ for any } x \text{ near } 0.$$

Therefore, we have functions $U(x) = |x|$ and $: (x) = -|x|$ such that

$$\lim_{x \to 0} U(x) = \lim_{x \to 0} L(x) = 0$$

which bound $f(x)$ near 0. Therefore, by Squeeze Theorem, we now know that

$$\lim_{x \to 0} x \sin\left(\frac{1}{x^2}\right) = 0.$$

Another useful theorem involving limits and continuity is the Intermediate Value Theorem!

**Theorem 2.21** (Intermediate Value Theorem). *Suppose $f$ is continuous on an interval $[a, b]$ and $M$ is a number between $f(a)$ and $f(b)$, then there is some point $c$ between $a$ and $b$ such that $f(c) = M$.*
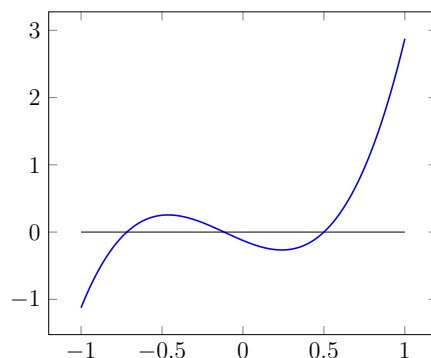
Figure 7: Graph of $f(x) = 3x^3 + x^2 - x - \frac{1}{8}$ on $[0, 1]$

The intuition behind this relies on the fact that if a function is continuous, then it has no 'hops,skips, breaks, or jumps'. Therefore, for a continuous function to go from greater than $M$ to less than $M$, the function must pass through $M$ at some point!

We can use the Intermediate Value Theorem to show the existence of roots for a given function.

**Example 2.22.** Show that $f(x) = 3x^3 + x^2 - x - \frac{1}{8}$ has a root.

Let's consider $f$ on the interval $[-1, 1]$. We can compute that $f(-1) = -1 - \frac{1}{8}$ and $f(1) = 3 - \frac{1}{8}$. We notice that

$$f(-1) < 0 < f(1).$$

Since $f(x)$ is continuous, we can apply the Intermediate Value Theorem which tells us that there is some point $c$ between 0 and 1 such that $f(c) = 0$. This, however, doesn't answer the question of how many roots $f(x)$ has in $[0, 1]$. The Intermediate Value Theorem guarantees that there is at least one root of $f$. If we graph $f(x)$, we see that it actually has 3 roots in $[0, 1]$.

We now move onto the actual definition of differentiability, so that we can answer our question about how to find the rate of change of a function at a given point.

# 3.   Differentiation

**Definition 3.1** (Definition of the Derivative)**.** We say that a function $f$ is *differentiable at $a$* if

$$f'(a) := \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \text{ exists.} \tag{3.2}$$

We call $f'(a)$ the derivative of $f(x)$ at $a$. If $f$ is differentiable at every point it is defined, we say that $f$ is differentiable and call $f'$ the derivative of $f$.

**Remark 3.3.** *Alternatively, we may write the derivative of $f$ as $\frac{df}{dx}(a)$ or $\frac{d}{dx}\left(f(x)\right)$. Here, $\frac{d}{dx}$ tells us to take the derivative of $f$.*

Let's unpack this definition. Recall that we can write the formula for a line between two points $(x_1, y_1)$ and $(x_2, y_2)$ as

$$y_2 - y_1 = m(x_2 - x_1)$$

where $m$ is the slope. In this case, we can calculate the slope of this line as

$$m = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}$$

which is the change in $y$ divided by the change in $x$. Substituting this into the point-slope formula, we see that

$$y_2 - y_1 = \underbrace{\left(\frac{\Delta y}{\Delta x}\right)}_{m}(x_2 - x_1).$$

Now looking at Definition 3.1, we see that the quotient

$$\frac{f(a+h) - f(a)}{h}$$

describes the slope of a line between the points $(a, f(a))$ and $(a+h, f(a+h))$, and we want to see what happens to this quotient as $h$ approaches 0.
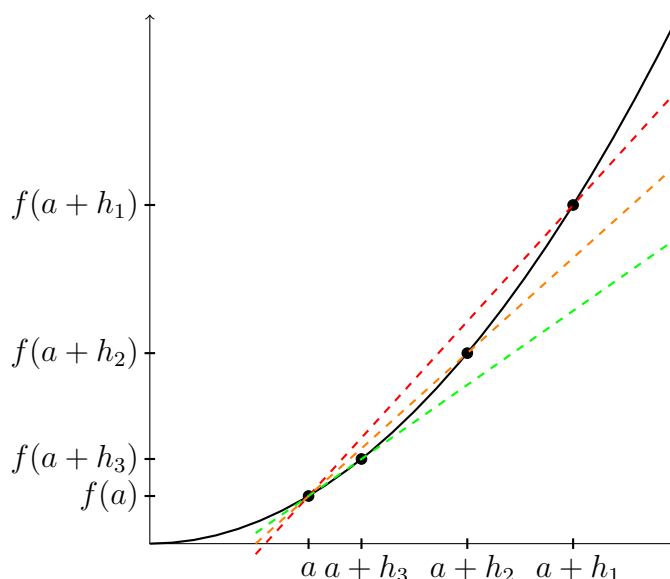
The picture above shows that as $h$ decreases, $\frac{f(a+h)-f(a)}{h}$ begins to resemble the slope of the line tangent to $f(x)$ at $a$. This shows us that the derivative gives us information about the *rate of change* of a function.

Motivated by the above picture, we see that if we let $x = a + h$, then our definition of the derivative is equivalent to the following one.

**Definition 3.4** (Alternative Definition of the Derivative). A function $f(x)$ is differentiable at $a$ if

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} \text{ exists.} \tag{3.5}$$

As before, we denote the value of this limit by $f'(a)$ or $\frac{df}{dx}(a)$.

**Remark 3.6.** *Borrowing the notation above, we can write $\Delta f = f(x) - f(a)$ and $\Delta x = x - a$, so that* (3.5) *above becomes*

$$\lim_{x \to a} \frac{\Delta f}{\Delta x} = \frac{df}{dx}(a).$$

Instead of taking the derivative at every single point $a$ that we're interested in, we can try to write the derivative of a function $f(x)$ as another function instead of fixing some point $a$! We can try an example of this.

**Example 3.7.** From algebra and geometry, we know that a circle with radius $r$ has area $\pi r^2$. What would we do if we wanted to know how quickly the area of the circle would change if we changed our radius. In this case, we write the area function as $A(r) = \pi r^2$. This time we use the definition of the derivative (Definition 3.1) using a variable $r$ for our radius instead of a single point $a$. Therefore, we can attempt to write the derivative $A'(r)$ as

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

Using our familiar algebra rules, we can show that $A(r+h) = \pi(r+h)^2 = \pi(r^2 + 2rh + h^2) = \pi r^2 + 2\pi rh + \pi h^2$. Therefore, we can rewrite this limit as

$$A'(x) = \lim_{h \to 0} \frac{\pi r^2 + 2\pi rh + \pi h^2 - \pi r^2}{h} = \lim_{h \to 0} \frac{2\pi rh + \pi h^2}{h} \tag{3.8}$$

$$= \lim_{h \to 0} \frac{h(2\pi r + \pi h)}{h} = \lim_{h \to 0} 2\pi r + \pi h = 2\pi r. \tag{3.9}$$

This shows us that the area of a circle of radius $r$ changes at a rate $2\pi r$. We can see this idea geometrically in () One thing that should stand out is that this is the same formula as the circumference $C(r)$ of the circle! Therefore, we showed that

$$\frac{dA}{dr} = C(r).$$

In words, this means that the derivative of the area of a circle with radius $r$ is its circumference! We probably wouldn't have noticed this pattern if we tried to find this derivative for specific points one by one.

     This technique is useful in a lot more cases, so we'll try to generalize this idea a bit. If we have a function $f(x)$ and can find another function that gives us the derivative of $f$ at any point $x$, we simply call it *the derivative of $f$* (without mentioning a specific point). Keeping our notation consistent, we'll write this function as $f'(x)$ or $\frac{df}{dx}$! This can be a little confusing at first, but the point is that the derivative of $f$ at a point $a$ is a number $f'(a)$, and the derivative of $f$ is another function $f'$.

     When we write it in this way, we can find that the derivative as we defined it has some very nice properties which will make the actual process computing derivatives much easier. You'll notice that several of the properties shown below resemble properties of limits which makes sense since we defined the derivative in terms of limits. These properties also harken back to the limit properties of continuous functions and later, we'll see that all differentiable functions are also continuous.

## 3.1   Differentiation Rules

As one might expect, using the definition of the derivative on each and every problem we encounter is tedious, boring, and generally uneventful. In order to get around this, we'll develop some tricks for computing derivatives easily using the derivatives of more 'elementary' functions.

**Proposition 3.10** (Scalar multiplication and Sum Rules)**.** *If $f$ and $g$ are differentiable at $a$ and $c$ is any real number, then $f + g$ and $c \cdot f$ are differentiable at $a$. Moreover,*

$$(f + g)'(a) = f'(a) + g'(a) \ and \tag{3.11}$$
$$(c \cdot f)'(a) = c \cdot f'(a). \tag{3.12}$$

*Proof.* We start with the definition of the derivative and through a short computation see that

$$\lim_{x\to a}\frac{(f+g)(x)-(f+g)(a)}{x-a}=\lim_{x\to a}\frac{f(x)+g(x)-f(a)-g(a)}{x-a}$$
$$=\lim_{x\to a}\frac{f(x)-f(a)+g(x)-g(a)}{x-a}$$
$$=\lim_{x\to a}\frac{f(x)-f(a)}{x-a}+\lim_{x\to a}\frac{g(x)-g(a)}{x-a}$$
$$=f'(a)+g'(a).$$

Therefore the derivative $(f+g)'(a)$ exists and equals $f'(a)+g'(a)$. A similar computation shows that $(cf)'(a)=c\cdot f'(a)$. $\qquad\square$

**Theorem 3.13** (Product Rule)**.** *If $f$ and $g$ are differentiable at $a$, then $f\cdot g$ is differentiable at $a$ and*

$$(f\cdot g)'=f'(a)\cdot g(a)+f(a)\cdot g'(a). \tag{3.14}$$

**Theorem 3.15** (Reciprocal Rule)**.** *If $f(x)$ is differentiable at $a$ and $f(a)\neq 0$, then $g(x)=\frac{1}{f(x)}$ is differentiable at $a$ and*

$$g'(a)=-\frac{f'(a)}{f(a)^2}. \tag{3.16}$$

**Corollary 3.17** (Quotient Rule)**.** *If $f$ and $g$ are differentiable at $a$ and $g(a)\neq 0$, then $Q(x)=\frac{f(x)}{g(x)}$ is differentiable $a$ and*

$$Q'(a)=\frac{f'(a)g(a)-f(a)g'(a)}{g(a)^2}. \tag{3.18}$$

Now that we have all these rules of differentiation, let's see how we can find the derivatives of a common class of functions.

**Question 3.19.** *What is the derivative of a polynomial function*

$$f(x)=a_n x^n+a_{n-1}x^{n-1}+\cdots+a_1 x+a_0, \tag{3.20}$$

*where $n=$ is a positive integer and each $a_i$ is any number?*

Let's start with the simple case where our function is constant.

**Proposition 3.21.** *The derivative of a constant function $f(x)=c$ is 0.*

*Proof.* If $f$ is constant, then $f(x+h)=f(x)=c$. Therefore,

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

$\qquad\square$

Moving onto the non-constant polynomials, the scalar multiplication and sum rules from above () tell us that we only need to be able to find the derivative of each $x^k$ to find the derivative of $f(x)$.

**Proposition 3.22** (Power Rule). *If $f(x) = x^n$ for any non-zero number $n$, then the derivative of $f$ is $f'(x) = nx^{n-1}$.*

*Proof.* We start with the definition of the derivative and substitute in our function $f$ i.e.

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{(x+h)^n - x^n}{h}. \tag{3.23}$$

Using the fact that

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k = a^n + na^{n-1}b + \frac{n(n-1)}{2!} a^{n-2}b^2 + \cdots + b^n, \tag{3.24}$$

where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, we calculate the limit as follows:

$$\lim_{h \to 0} \frac{(x+h)^n - x^n}{h} = \lim_{h \to 0} \frac{(x^n + nx^{n-1}h + \frac{n(n-1)}{2!} x^{n-2}h^2 + \cdots + h^n) - x^n}{h} \tag{3.25}$$

$$= \lim_{h \to 0} \frac{nx^{n-1}h + \frac{n(n-1)}{2!} x^{n-2}h^2 + \cdots + h^n}{h} \tag{3.26}$$

$$= \lim_{h \to 0} \left[ nx^{n-1} + \frac{n(n-1)}{2!} x^{n-2}h + \cdots + h^{n-1} \right]. \tag{3.27}$$

Using the additive and multiplicative properties of limits, we see that

$$f'(x) = \lim_{h \to 0} \left[ nx^{n-1} + \frac{n(n-1)}{2!} x^{n-2}h + \cdots + h^{n-1} \right] = nx^{n-1}, \tag{3.28}$$

which proves the theorem. $\square$

**Example 3.29.** Using the power rule, find the derivative of $f(x) = x^2 + 3x$.

There are other useful derivative formulas that one should know when doing and learning calculus.

**Proposition 3.30.** *The functions $\sin(x), \cos(x)$ and $e^x$ are differentiable.*

1. *The derivative of $\sin(x)$ is $\cos(x)$,*

2. *the derivative of $\cos(x)$ is $-\sin(x)$,*

3. *the derivative of $e^x$ is $e^x$, and*

4. *the derivative of $\ln(x)$ is $\frac{1}{x}$.*

The proofs for the first two parts rely on the following additive identities of trigonometry:

$$\sin(a + b) = \sin(a)\cos(b) + \cos(a)\sin(b), \tag{3.31}$$
$$\cos(a + b) = \cos(a)\cos(b) - \sin(a)\sin(b), \tag{3.32}$$

and using the following limit identities,

$$\lim_{x \to 0} \frac{\cos(x) - 1}{x} = 0 \text{ and } \lim_{x \to 0} \frac{\sin(x)}{x} = 1. \tag{3.33}$$

We'll leave the proof of the first two as an exercise, and we'll prove the third and fourth later when we do implicit differentiation.

Now that we understand how to differentiate function like $x^2$ or $\sin(t)$ as well as their sums, products, and quotients, we can begin to tackle a number of problems. But we've yet to answer how we can differentiate *compositions* of functions like $\sin(t^2)$ or $\cos^3(t) + e^{2t}$. That is, we want to understand how we can differentiate functions within other functions. The first step of this is identifying what is a composition of functions. Say we're given the function $\sin(t^2)$ as above. We can clearly see that this is made up of $t^2$ and the familiar $\sin(x)$. To clarify this notation, we write $f(x) = \sin(x)$ and $g(t) = t^2$. This way, we can see that choosing $x = g(t)$ allows us to write

$$f(x) = f(g(t)) = f(t^2) = \sin(t^2).$$

This still doesn't answer the question of differentiation. Let's set up this question more abstractly. Say we have any function $f(x)$ and another function $g(t)$, then our goal is to find the derivative of its composition,

$$(f \circ g)'(t) = \frac{df}{dt}.$$

Going back to most basic idea behind the derivative, we know that $f'(x) = \frac{df}{dx}$ tells us the rate of change of $f$ as we move near $x$. Therefore, the derivative $f'(g(t))$ should tell us the rate of change near $g(t)$, but since $g$ itself is a function, we must account for the change in $g(t)$ near $t$ i.e. $\frac{dg}{dt}$. Since $g$ is acting as our input $x$, we can write its derivative as $\frac{dx}{dt}$.

DONT DO THIS YOURseLF

$$\frac{df}{dt} = \frac{df}{dx}\frac{dx}{dt}.$$

Handwavely, but the equation is ...

Write as proposition.

**Proposition 3.34** (The Chain Rule)**.** *If $f$ is differentiable at $g(t)$ and $g$ is differentible at $t$, then the derivative of $f \circ g$ is*

$$(f \circ g)'(t) = g'(t) \cdot f'(g(t)).$$

**Remark 3.35.** *I often remember this as 'the derivative of the inside function times the outside function'.*

**Examples 3.36.** Compute the derivative $f'(x)$ for the following functions:

**Example A.1.** $f(x) = x^5 + 2x + 1$,

**Example A.2.** $f(x) = \sin^3(x) = \big(\sin(x)\big)^3$,

**Example A.3.** $f(x) = x^3 \sin(x)$,

**Example A.4.** $f(x) = \frac{5x}{x^2+1}$,

**Example A.5.** $f(x) = e^{6x}$

**Example A.6.** $f(x) = x^2 e^{2x}$

**Example A.7.** $f(x) = \frac{1}{x^2}$

**Example A.8.** $f(x) = e^{-6x} \ln(x)$

**Example A.1.**
Since we're differentiating a sum, we can take the derivative of each of its parts! Using the power rule (), we see that the derivative of $x^5$ is $5x^4$, the derivative of $2x$ is 2 and the derivative of 1 is 0. This means that the derivative of $x^5 + 2x + 1$ is given by $5x^4 + 2$.

**Example A.2.**
This problem is an application of the chain rule (add citation) and the power rule! Using the $\sin(x)$ as the 'inside' function $(g(x))$ and $x^3$ as the 'outside' function $(f(x))$. Starting with the 'outside' function $f$ we use the power rule to show

$$f(x) = x^3 \xrightarrow{\frac{d}{dx}} f'(x) = 3x^2.$$

by the power rule. Similarly, the derivative of $g(x) = \sin(x)$ is $g'(x) = \cos(x)$ by Proposition 3.30. Following the chain rule formula, we can compute $f'(g(x)) = 3\sin^2(x)$. Therefore, we write the derivative as $3\cos(x)\sin^2(x)$.

**Example A.3.**
**Example A.4.**
**Example A.5.**
**Example A.6.**
**Example A.7.**
**Example A.8.**

The derivative is also useful for answering questions about the average rate of change as well as the maxima and minima.

## 3.2    Differentiation Theorems

**Theorem 3.37** (Rolle's Theorem). *Suppose $f$ is continuous on a closed interval $[a, b]$ and differentiable on the open interval $(a, b)$ and $f(a) = f(b)$, then there is some point $c \in [a, b]$ such that $f'(c) = 0$.*

With Rolle's theorem, we can define the following theorem which allows us to relate the secant lines discussed earlier with the derivative.

**Theorem 3.38** (Mean Value Theorem). *Suppose $f$ is continuous on a closed interval $[a, b]$ and differentiable on the open interval $(a, b)$, then there is some point $c$ in $[a, b]$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \tag{3.39}$$

*Proof.* Let $g(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x - a)$. The function $g$ is differentiable with $g(a) = f(a)$ and $g(b) = f(b)$. Using differentiation rules described earlier,

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}. \tag{3.40}$$

Therefore, we can apply Rolle's theorem and see that there is some $c$ such that $g'(c) = 0$, so

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \tag{3.41}$$

$\square$

Essentially, this theorem states that the secant line of $f$ at the points $a$ and $b$ has the same slope as the derivative $f$ at some point $c$ or *the average rate of change on $[a, b]$ is realized as the instantaneous rate of change at a point $c$ in $[a, b]$.*

# 4.   Applications of Differentiation

## 4.1   Extrema: Maxima and Minima

This leads us to a very important application of calculus showing the existence of and finding extrema. That is, finding the points at which functions attain their extrema i.e. their maxima or minima.

**Definition 4.1.** We say that $f$ has a *local maximum* at $c$ if

$$f(c) \geq f(x) \text{ for all } x \text{ sufficiently close to } c.$$

Likewise, we define the *global maximum of $f$* to be the point $c$ such that

$$f(c) \geq f(x) \text{ for all } x \text{ where } f \text{ is defined.}$$

We can define local and global minima similarly by flipping the inequality.

One thing to notice is that not all functions have local extrema. For example, take the function

$$f(x) = x \text{ where } 0 < x < 1$$

or

$$f(x) = \begin{cases} 0 \text{ when } x \text{ is in } \mathbb{Q}, \\ 1 \text{ otherwise,} \end{cases}$$

which has no local maxima or minima though every point is an absolute maximum or minimum. Luckily, the Extreme Value Theorem gives us criteria to tell when a function has absolute extrema.

**Theorem 4.2** (Extreme Value Theorem)**.** *If $f$ is continuous on a closed interval $[a, b]$, then $f$ has both a global maximum and minimum on that interval.*

This gives us a neat answer in the case $f$ is continuous on a closed interval $[a, b]$, but what happens if we further assume that $f$ is differentiable on the open interval $(a, b)$.

Let's consider the function $f(x) = 1 - x^2$ on the interval $[-1, 1]$. Clearly, $f$ is continuous on $[-1, 1]$ and differentiable on $(-1, 1)$. By the Extreme Value Theorem, you can see that we're guaranteed a maximum and minimum on $[-1, 1]$. Graphing the function, we see that the minimum occurs at $x = 1, -1$ where $f(x) = 0$ and that the maximum occurs at $x = 0$ where $f(x) = 1$. Let's see what this entails for the derivative of $f$.

The derivative of $f(x)$ is $f'(x) = -2x$. From this, we can notice that if $x < 0$, then $f'(x) > 0$. That is $f(x)$ has a positive slope to its tangent and is *increasing*. Similarly, when $x > 0$, $f'(x) < 0$ and $f$ is *decreasing*. Furthermore, $f'(0) = 0$ which makes sense since for $f$ to achieve a maximum at 0, it must increase to a point and then decrease from there on for

$x$ near 0. This means that not only does $f$ have a global maximum at 0; it also has a local maximum.

Above we used the sign of the derivative $f'$ to determine whether $f$ was increasing or decreasing. In general, this holds for every differentiable function $f$.

**Proposition 4.3.** *Given a differentiable function $f$,*

1. *If $f'(x) > 0$, then $f$ is increasing at $x$.*

2. *If $f'(x) > 0$, then $f$ is decreasing at $x$.*

This leaves us to understand the case where $f'(x) = 0$. We can begin to describe this case using what is known as the *first derivative test.*

**Theorem 4.4** (First Derivative Test)**.**

## 4.2 Optimization

## 4.3 Linear Approximations

By construction, the derivative describes the instantaneous rate of change of a function at a particular point. As we saw, this is the slope of the tangent line at that point. Because of this, the derivative can be used to provide an estimate of what the function is doing near a given points. That is, if we fix a function $f$ and a point $a$ where $f$ is differentible, we can use the derivative $f'(a)$ and its corresponding tangent line to $f$ to approximate the function $f$ at points $x$ near $a$. This process is called *linear approximation.*

## 4.4 Implicit Differentiation

In the cases given previously, we have always been able to isolate our expressions as functions of single variable i.e

$$f(x) = x^2 + 5x - 3 \text{ or } y = \frac{\sin(x+1)}{x^2}.$$

This may not always be possible or the simplest way to do things! Suppose we're given the equation for a circle centered at the origin with radius 1

$$x^2 + y^2 = 1.$$

Yes, it is possible to isolate $y$ and find $y = \pm\sqrt{1 - x^2}$, but this leaves two different cases to watch out for when finding the derivative on the circle.

Instead, if we remember that $y$ is a function as $x$, we can differentiate this equation without isolating $y$ using the *chain rule* as follows:

$$\frac{d}{dx}\left[x^2 + y^2 = 1\right] \implies 2x + 2y\frac{dy}{dx} = 0. \tag{4.5}$$
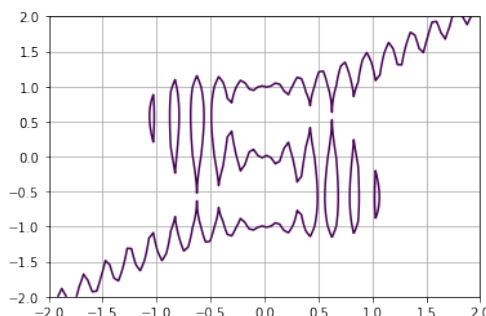
Figure 8: Graph of (4.7)

Rearranging this, we see that

$$\frac{dy}{dx} = -\frac{x}{y} \text{ where } x, y \text{ satisfy the equation we started with.}$$

This technique is called *implicit differentiation* and comes in handy when dealing with complex expressions in terms of both $y$ and $x$. Let's do another example.

**Example 4.6.** Let's consider the curve given by

$$y^3 - y = x^3 - x\cos(30x). \tag{4.7}$$

Suppose we wanted to find the slope of a point tangent to this curve. Instead of hopelessly trying to simplify this equation to solve for $y$ and differentiate, we can differentiate implicitly. Differentiating (4.7) with respect to $x$, we get

$$3y^2\frac{dy}{dx} - \frac{dy}{dx} = 3x^2 - \cos(30x) - 30x\sin(30x).$$

Pulling out the $\frac{dy}{dx}$, we see

$$\frac{dy}{dx} = \frac{3x^2 - \cos(30x) + 30x\sin(30x)}{3y^2 - 1}.$$

Notice that this derivative is undefined along the lines $y = \pm\frac{1}{\sqrt{3}}$, which correspond the points where the line tangent to the graph is vertical. Similarly, we can find the points where the slope is horizontal by finding the zeroes of $3x^2 - \cos(30x) + 30x\sin(30x)$.

We can can use implicit differentiation to prove Proposition 3.30.3. using Proposition 3.30.4. i.e. that derivative of $f(x) = e^x$ is $e^x$.

*Proof.* Suppose we have $y = e^x$. This is equivalent to saying that $\ln y = x$. Taking the derivative of both sides with respect to $x$, we see that

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

using that $\frac{d}{dy}(\ln y) = \frac{1}{y}$. Therefoere, $\frac{dy}{dx} = y = e^x$. $\qquad\square$

**Exercise 4.8.** Prove that $\frac{d}{dx}(\ln x) = \frac{1}{x}$ using that the derivative of $e^x$ is itself.

## 4.5   Related Rates

# 5.   Integration

We'll now begin to develop the concept of *integration*. Where differentiation was concerned with the rate of change at a particular point, integration is concerned with the way these changes accumulate over a range of points. In order to understand this mathematically, we'll need to develop some convenient notion for dealing with sums with a large number of terms. To do this, we'll rely on Riemann sums, but first we'll introduce the "capital sigma" notation for sums for those who are unfamiliar.

Say we have numbers $a_1, a_2, a_3, a_4$, we can write their sum as $a_1 + a_2 + a_3 + a_4$ or as $\sum_{i=1}^{4} a_i$ where $i$ is called the 'dummy variable'. This is used to make the notation simpler and it'll reoccur often when have many terms to add together. More generally, we can write the sum of numbers $a_1, a_2, \ldots, a_n$ as

$$\sum_{i=1}^{n} a_i := a_1 + a_2 + \cdots + a_n. \tag{5.1}$$

One thing to notice is that the sum is *linear*. This property is key to what follows.

**Lemma 5.2** (Linearity of the sum)**.** *For any numbers $a_1, \ldots a_n$ and $b_1, \ldots, b_n$ and constant c, we have:*

1. $\sum_{i=1}^{n} a_i = \sum_{i=1}^{k} a_i + \sum_{i=k+1}^{n}$ *for any $k$ between $1$ and $n$,*

2. $\sum_{i=1}^{n} (a_i + b_i) = \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i,$

3. $\sum_{i=1}^{n} (ca_i) = c \sum_{i=1}^{n} a_i.$

This linearity also allows us to establish to the following important identities.

**Lemma 5.3** (Formulas for Integer Sums)**.** *The following identities hold.*

1. $\sum_{i=1}^{n} i = \frac{n(n+1)}{2},$

2. $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6},$

3. $\sum_{i=1}^{n} i^3 = \left( \frac{n(n+1)}{2} \right)^2.$

*Proof.* These can be proved by induction using the linearity of the sum as established in Lemma 5.2. $\qquad\square$

## 5.1   Riemann Sums

These identities will be helpful in establishing the integrals in several of the sums that will appear. When we "integration is concerned with the way changes accumulate", we often want to this as the integral as the area under the curve of a function over an interval.
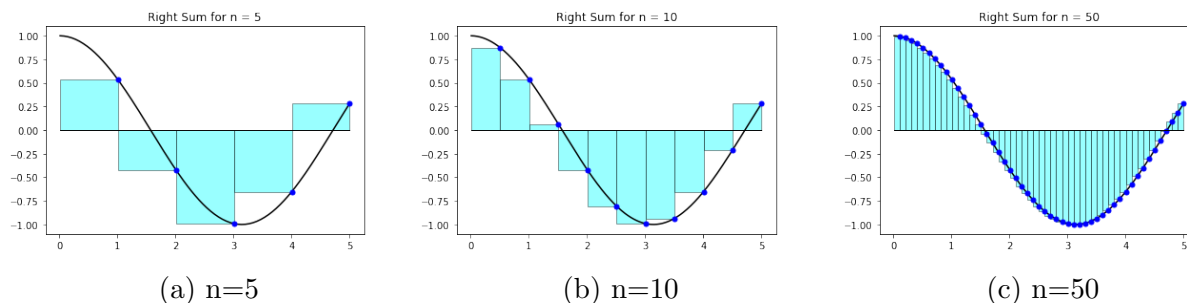
(a) n=5           (b) n=10           (c) n=50

Figure 9: Upper Riemann Sums for $\cos(x)$ on $[0, 5]$ using $n$ subintervals of equal length.

Therefore, we start by trying to approximate this area under the curve using rectangles as in Fig. 9. In order to this, we have to divide our interval into pieces that will from the bases of these rectangles. This is the idea of the partition.

**Definition 5.4** (Partition). If we have an interval $[a, b]$, we define a *partition $P$ of $[a, b]$* to be the set of points $P = \{x_0, x_1, \ldots, x_n\}$. Additionally, we define the *mesh* of the partition $||P||$ to be the maximum length of the intervals $[x_i, x_{i+1}]$ in $P$.

With partitions under our belt, we can define the Riemann sum: our approximation of the area underneath the curve given a particular partition.

**Definition 5.5** (Definition of the Riemann Sum). Given a function $f$ on the interval $[a, b]$ and a partition $P = \{x_0, x_1, \ldots, x_n\}$ of [a,b], we define the *upper Riemann sum of $f$ with respect to $P$* as

$$U(f; P) := \sum_{i=1}^{n} f(x_i)(x_i - x_{i-1}). \tag{5.6}$$

Similarly, we can define the *lower Riemann sum* as

$$L(f; P) := \sum_{i=1}^{n} f(x_{i-1})(x_i - x_{i-1}). \tag{5.7}$$

Between this two, we can also define the *midpoint Riemann sum* by taking the midpoints of each interval, so that

$$M(f; P) := \sum_{i=1}^{n} f\left(\frac{x_i + x_{i-1}}{2}\right)(x_i - x_{i-1}). \tag{5.8}$$

Notice each term in this sum is of the form $f(x)(x_i - x_{i-1})$, this is the area of a rectangle with height $f(x)$ and width $(x_i - x_{i-1})$. In the case of the upper sum, we take the value of $f$ at rightmost end point of the interval $[x_{i-1}, x_i]$ with the left most value of the interval being used for the lower sum. If we want a function to have a definite area under it, we'd expect that the upper and lower sums fo become closer and closer as the mesh of our partition decreases since the left endpoint will approach the right endpoint of each interval. Since we'd want the intervals to be 'infinitely' thin to ensure accuracy, this motivates us to define the integral in terms of limits.

## 5.2    Definite Integrals

**Definition 5.9** (Definition of the Definite Integral)**.** We define the *(definite) integral of $f$ over $[a, b]$* to be

$$\int_a^b f(x)dx := \lim_{||P|| \to 0} U(f; P) = \lim_{||P|| \to 0} L(f; P) \tag{5.10}$$

provided both limits exist and are equal. That is, the integral is the defined as the limit of our Riemann sum as the number of rectangles goes to infinity and each becomes infinitely thin having infinitesimal width $dx$.

**Remark 5.11.** *When the integral as defined above exists, we say that $f$ is* integrable *on $[a, b]$.*

This raises the question of what kinds of functions are integrable for this. To provide a partial answer, we have the following proposition.

**Proposition 5.12.** *When the function $f$ is continuous on $[a, b]$, the integral $\int_a^b f(x)dx$ exists.*

Though we know that every continuous function is integrable, we also know that not every function is integrable.

Notice that the limit in Definition 5.9 is independent of partition so that if we have a sequence of partitions $P_n$ whose mesh go to 0, then we can compute the integral of $f$, so that

$$\int_a^b f(x)dx = \lim_{n \to \infty} U(f; P_n) = \lim_{n \to \infty} L(f; P_n). \tag{5.13}$$

**Example 5.14.** Suppose we want to calculate the integral of $f(x) = x^2$ on $[0, 1]$. By crefIntExists, we can know $\int_0^1 x^2 dx$ exists and pick sequence of partitions $P_n$ where each partition is given by

$$P_n = \left\{ 0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1 \right\} = \left\{ \frac{i}{n} \mid 0 \leq i \leq n \right\}, \tag{5.15}$$

dividing $[0, 1]$ into $n$ subintervals of equal length. Since $||P||$ is clearly $\frac{1}{n}$, we can write the upper Riemann sum as

$$U(f; P_n) = \sum_{i=1}^n f\left(\frac{i}{n}\right)\left(\frac{i}{n} - \frac{i-1}{n}\right) = \sum_{i=1}^n \frac{i^2}{n^2}\left(\frac{1}{n}\right) = \frac{1}{n^3}\sum_{i=1}^n i^2. \tag{5.16}$$

Using Lemma 5.3 we can reduce the above equations to

$$U(f; P_n) = \frac{n(n+1)(2n+1)}{6n^3} = \frac{1}{3} + \frac{3}{n} + \frac{1}{n^2}. \tag{5.17}$$

Taking the limit, we see that $\int_0^1 x^2 dx = \frac{1}{3}$ by (5.13).

---

This process can be visualized using computer code as in Fig. 9. Though we can use Riemann sums to solve for the integrals of simple functions directly, this method is tedious and doesn't always come easily. Therefore, we will need to develop some method of easily computing integrals, but first we must establish some more properties of the integral.

**Lemma 5.18** (Linearity of the Integral). *The integral is linear in the following sense: For any integrable $f$ and $g$ on $[a, b]$,*

1. *$\int_a^c f(x)dx + \int_c^b f(x)dx$ if $a < c < b$,*

2. *$\int_a^b [f(x) + g(x)]dx = \int_a^b f(x)dx + \int_a^b g(x)dx,$*

3. *$\int_a^b cf(x)dx = c\int_a^b f(x)dx$ for any constant $c$.*

4. *$\int_a^a f(x)dx = 0$*

5. *$\int_a^b f(x)dx = -\int_b^a f(x)dx.$*

**Remark 5.19.** *Notice that the first three properties mirror those of the sum that we presented in Lemma 5.2.*

A consequence of this similarity is that the sum *commutes* with the integral.

**Corollary 5.20.** *If $f_1, f_2, \ldots, f_n$ are integrable functions on $[a, b]$, then*

$$\int_a^b \left( \sum_{i=1}^n f_i(x) \right) dx = \sum_{i=1}^n \left( \int_a^b f_i(x)dx \right). \tag{5.21}$$

We can restate Corollary 5.20 as "the sum of the integral is the integral of the sum". Furthermore, we can deduce more properties of the integral such the integral of a non-negative function must be non-negative. With this in mind, we can get a very crude bound on the definite integral of a function.

**Proposition 5.22.** *If $f$ is an integrable function on $[a, b]$ and $m$ and $M$ the minimum and maximum on $[a, b]$ respectively, then*

$$m(b - a) \le \int_a^b f(x)dx \le M(b - a). \tag{5.23}$$

**Remark 5.24.** *This immediately tells us that for any constant function $c$ on $[a, b]$, $\int_a^b c\, dx = c(b - a)$.*

# 6. Fundamental Theorems of Calculus

## 6.1 Indefinite Integrals

In general, we want to be able to compute the actual value of the integral in an easy way, not just establish a crude estimate or use Riemann sums to compute the value as it can be tedious or impossible. In order to make the process of finding integrals easier, we'll begin to establish a relationship between differentiation and integration. We'll start with the notion of an anti-derivative.

**Definition 6.1** (Anti-derivative). If $f$ is a function on $[a, b]$, we say that a differentiable function $F$ is an *anti-derivative* if $F'(x) = f(x)$ for all $x \in [a, b]$.

**Proposition 6.2.** *If $F$ is an anti-derivative of $f$, then $F(x) + C$ is also an anti-derivative of $f$ for any constant $C$. Moreover, if $F_1$ and $F_2$ are anti-derivatives, then they differ by a constant.*

*Proof.* Since the derivative of any constant $C$ is 0, the derivative of $F$ must equal the derivative of $F(x) + C$. Similarly, if $F_1$ and $F_2$ are two anti-derivatives of $f$, then $F_1 - F_2$ must have derivative 0. Therefore, their difference is constant, so that $F_1(x) = F_2(x) + C$ for some $C$. $\square$

**Definition 6.3** (Indefinite Integral). With Proposition 6.2 in mind, we define the *indefinite integral of $f$* as

$$\int f(x)dx := F(x) + C, \tag{6.4}$$

where $F$ is an anti-derivative of $f$. We use $C$ as placeholder constant to denote all positive anti-derivatives of f.

Armed with this, we can now state one of the two Fundamental theorems of calculus.

**Theorem 6.5** (The First Fundamental Theorem of Calculus). *Suppose $f$ is a continuous function on the interval $[a, b]$. If we define*

$$F(x) = \int_a^x f(t)dt, \text{ then } F'(x) = f(x), \tag{6.6}$$

*i.e. $\int_a^x f(t)dt$ is an anti-derivative of $f(x)$.*

Therefore, any continuous function of an interval has an anti-derivative on that interval.

**Theorem 6.7** (The Second Fundamental Theorem of Calculus). *If $f$ is integrable on $[a, b]$, then*

$$\int_a^b f(x)dx = F(b) - F(a) \tag{6.8}$$

*for any anti-derivative $F$.*

**Remark 6.9.** *Notice that the Second Fundamental Theorem of Calculus is independent of our choice of anti-derivative by Proposition 6.2, so we can always set $C = 0$. That is, provided $F$ exists. We will often use $F(x)|_a^b$ as shorthand for $F(b) - F(a)$, emphasizing the relationship between the indefinite and definite integrals.*

Theorem 6.5 and Theorem 6.7 are called the Fundamental Theorems of Calculus (FTC for short) because they establish the relationship between differentiation and integration. The idea being that the derivative and integration are functionally opposite operatives. One immediate result of the Theorem 6.7 is that we have drastically simplified the problem of computing the definite integral; we simply need to find an anti-derivative.

The Fundamental Theorems of Calculus make the problem of finding the definite integral of functions like $f(x) = x$ much simpler as we can compute that

$$\int_a^b x \, dx = \frac{b^2}{2} - \frac{a^2}{2} \tag{6.10}$$

since $\frac{x^2}{2}$ is an anti-derivative of $x$. In fact, we can easily derive a more general statement for the indefinite integral of powers of $x$.

**Proposition 6.11** (Reverse Power Rule). *If $f(x) = x^n$ for some $n \neq -1$, then $f$ has anti-derivative $F(x) = \frac{x^{n+1}}{n+1}$ i.e*

$$\int x^n \, dx = \frac{x^{n+1}}{n+1} + C. \tag{6.12}$$

**Remark 6.13.** *In particular, this shows that for any constant $c = cx^0$, we have $\int c \, dx = cx$ which is consistent with Remark 5.24.*

Using the linearity of the integral, we can use Proposition 6.11 to derive a formula for the indefinite integral (or anti-derivative) of any polynomial.

**Corollary 6.14.** *If we have a polynomial $p(x) = a_n x^n + a_{n-1} x^{n-1} \cdots + a_1 x + a_0$ where the $a_i$ are constants, then*

$$\int p(x)dx = \left(\frac{a_n}{n+1}\right) x^{n+1} + \left(\frac{a_{n-1}}{n}\right) x^n + \cdots + \left(\frac{a_1}{2}\right) x^2 + a_0 x + C. \tag{6.15}$$

*Proof.* Writing our polynomial as $p(x) = \sum_{i=0}^n a_i x^i$ using sum notation, we can use Corollary 5.20 to see that

$$\int p(x)dx = \int \left(\sum_{i=0}^n a_i x^i\right) dx = \sum_{i=0}^n \left(\int a_i x^i dx\right). \tag{6.16}$$

Therefore, we can use Proposition 6.11 to compute the integral on the righthand side, so that

$$\int p(x)dx = \sum_{i=0}^n \left(\frac{a_i}{i+1}\right) x^{i+1} + C. \tag{6.17}$$

$\square$

## 6.2 Consequences of the Fundamental Theorems

We can derive similar formulas for the integrals of the trigonometric functions with

$$\int \sin(x) \ dx = -\cos(x) + C \ \text{and} \ \int_a^b \cos(x) \ dx = \sin(x) + C, \tag{6.18}$$

which for now we will take as granted.

We can also define formulas for the integrals composition of functions.

**Theorem 6.19** (*u*-Substitution)**.** *For functions $f, u$ such that $u$ is differentiable on $[a, b]$ and $f$ is continuous on the image $u([a, b])$,*

$$\int_a^b f(u(x))u'(x)dx = \int_{u(a)}^{u(b)} f(x)dx. \tag{6.20}$$

*Proof.* Since $f$ is continuous, it has an anti derivative $F$. Noticing that $F \circ u$ has derivative $F'(u(x))u'(x)$ by chain rule, we see that $F \circ u$ is an anti derivative of $f(u(x))u'(x)$. Therefore, Theorem 6.7 tells us that

$$\int_a^b f(u(x))u'(x)dx = F(u(b)) - F(u(a)) = \int_{u(a)}^{u(b)} f(x)dx. \tag{6.21}$$

$\square$

**Remark 6.22.** *There is an direct analogue of this for indefinite integrals which we will use later.*

Notice this resembles the chain rule, but in reverse. This allows us to reduce integrals of many functions into the cases we've described above. Similarly, we can prove a sort of reverse product rule.

**Theorem 6.23** (Integration by Parts)**.** *For differentiable functions $u, v$ on $[a, b]$,*

$$\int_a^b u \ dv = (uv)|_a^b - \int_a^b v \ du. \tag{6.24}$$

*where $dv$ and $du$ denote $v'(x)dx$ and $u'(x)dx$ respectively.*

*Proof.* Consider the function $u \cdot v$. By the product rule, its derivative is $uv' + vu'$. Therefore, linearity of the integral and FTC tell us

$$(uv)|_a^b = \int_a^b [u(x)v'(x) + v(x)u'(x)]dx = \int_a^b u(x)v'(x)dx + \int_a^b v(x)u'(x)dx. \tag{6.25}$$

Making the substitutions $dv = v'(x)dx$, $du = u'(x)dx$ and subtracting over the second term, we get the result

$$\int_a^b u \ dv = (uv)|_a^b - \int_a^b v \ du.$$

$\square$

Furthermore, we can see that integral is consistent with our geometric intuition of area under the curve when it comes to functions with symmetry.

**Proposition 6.26** (Integrating Even and Odd Functions). *Suppose $f$ is continuous on $[-b, b]$, then*

$$\int_{-b}^{b} f(x)dx = 2 \int_{0}^{b} f(x)dx \text{ if } f \text{ is even, and} \tag{6.27}$$

$$\int_{-b}^{b} f(x)dx = 0 \text{ if } f \text{ is odd.} \tag{6.28}$$

**Proposition 6.29** (Integrating Periodic Functions). *If $f$ is continuous and periodic with period $T$, then*

$$\int_{x}^{x+T} f(x)dx = \int_{0}^{T} f(x)dx \text{ for any } x. \tag{6.30}$$

We can use these to compute the followings examples.

**Examples 6.31.** Compute the indefinite integral $\int f(x)dx$ for the following functions:

**Example A.1.** $f(x) = x^5 + 2x + 1,$

**Example A.2.** $f(x) = \cos(x)\sin^2(x),$

**Example A.3.** $f(x) = \sin^2(x) + \cos^2(x),$

**Example A.4.** $f(x) = \frac{3x}{\sqrt{x^2+1}},$

**Example A.5.** $f(x) = 3x\cos(x)$

Starting with **Example A.1.**, we can see that this is an application of Corollary 6.14, so that

$$\int (x^5 + 2x + 1) \, dx = \frac{x^6}{6} + x^2 + x + C.$$

For **Example A.2.**, we'll need $u$-substitution. Noticing that the derivative of $\sin^2(x)$ will have a $\cos(x)$ in front by Chain Rule ($\frac{d}{dx}(\sin x) = \cos x$), we define $u = \sin(x)$ so that $du = \cos(x)dx$. Therefore, we can make these substitutions and see

$$\int \cos(x)\sin^2(x)dx = \int \underbrace{\sin^2(x)}_{u^2}\underbrace{\cos(x)dx}_{du} = \int u^2 du = \frac{1}{3}u^3 + C.$$

To finish, we have to back substitute, so that everything is in terms of $x$, so our final answer is

$$\int \cos(x)\sin^2(x)dx = \frac{1}{3}\sin^3(x) + C.$$

**Example A.3.** comes down to the trig identity $\sin^2(x) + \cos^2(x) = 1$, so that

$$\int \sin^2(x) + \cos^2(x)dx = \int 1 \ dx = x + C.$$

Moving onto **Example A.4.**, we see that the function $x^2 + 1$ has derivative $2x$, so we can try $u$-substitution again. Making the substitution $u = x^2 + 1$, we get $du = 2x$, so that $\frac{3}{2}du = 3x \ dx$. Therefore, making these substitutions, we compute

$$\int \frac{3x}{\sqrt{x^2 + 1}} \ dx = \frac{3}{2} \int \frac{1}{\sqrt{u}}du = \frac{3}{2} \int u^{-\frac{1}{2}}du \tag{6.32}$$

$$= \frac{3}{2} \cdot 2u^{\frac{1}{2}} + C, \text{ by Proposition 6.11} \tag{6.33}$$

$$= 3(x^2 + 1)^{\frac{1}{2}} + C, \text{ back substituting } u = x^2 + 1. \tag{6.34}$$

We can see that **Example A.5.** is the product of two functions, neither of which contains the derivative of the other, so $u$-substitution will not work. Therefore, we try the product rule. Letting, $u = 3x$ and $dv' = \cos(x)$, we have $u' = 3$ and $v = \sin(x)$. Plugging this into the formula, we get

$$\int 3x \cos(x)dx = 3x \sin(x) - \int 3 \sin(x)dx = 3x \sin(x) - 3\cos(x) + C.$$

Let's try to do the same for definite integrals.

Compute the definite integral $\int_a^b f(x)$ for the following functions:

**Example B.1.** $f(x) = x^3, a = -2, b = 4$

**Example B.2.** $f(x) = \sin^{5001}(x), a = -1, b = 1,$

**Example B.3.** $f(x) = 4x(x^2 + 8)^3,$

**Example B.4.** $f(x) = \sin(x)\cos^2(x), a = -\frac{3\pi}{2}, b = \frac{5\pi}{2}$

**Example B.5.** $f(x) = |\cos(x)|, a = 0, b = 2000\pi$

Starting with **Example B.1.**, we can see this is just an application of the reverse power rule, so that

$$\int_{-2}^4 x^3 = \left(\frac{x^4}{4}\right)\Big|_{-2}^4 = \frac{4^4}{4} - \frac{(-2)^4}{4} = 64 - 4 = 60.$$

**Example B.2.** looks to be ridiculously hard to solve directly. We'll need to be clever here. Noticing that $\sin x$ is odd and that, therefore, $\sin^{5001}(x)$ must be odd, we can use this symmetry as in Proposition 6.26 to see that $\int_{-1}^1 \sin^{5001}(x)dx = 0$.

**Example B.3.** and **Example B.4.** are applications of $u$-substitution and will be left as exercises.

Lastly, **Example B.5.** can be solved using the periodic of $|\cos(x)|$. Since $|\cos(x)|$ has period $\pi$, we can decompose the integral as

$$\int_0^{2000\pi} |\cos(x)|\,dx = 2000 \int_0^{\pi} |\cos(x)|\,dx$$

using Proposition 6.29. This leaves us to calculate $|\cos(x)|$. Since $\cos(x)$ is non-negative on $[0, \frac{\pi}{2}]$ and negative on $[\frac{\pi}{2}, \pi]$, we can once again decompose the integral to find

$$\int_0^{\pi} |\cos(x)|\,dx = \int_0^{\frac{\pi}{2}} \cos(x)dx - \int_{\frac{\pi}{2}}^{\pi} \cos(x)dx = \sin(x)|_0^{\frac{\pi}{2}} - \sin(x)|_{\frac{\pi}{2}}^{\pi} = 2\sin\left(\frac{\pi}{2}\right) = 2.$$

Therefore, our final answer is

$$\int_0^{2000\pi} |\cos(x)|\,dx = 4000.$$

With our knowledge that differentiation and integration are a kind of reciprocal operations via the FTC, we can prove an analogue of the Mean Value theorem for integrals.

**Theorem 6.35** (Mean Value Theorem for Integrals). *If $f$ is continuous on $[a, b]$, then there exists some point $c$ in $[a, b]$ such that*

$$f(c) = \frac{1}{b-a} \int_a^b f(x)dx. \tag{6.36}$$

*Proof.* Since $f$ is continuous, we have an anti-derivative $F(x) = \int_a^x f(t)dt$ which is differentiable on $(a, b)$ and continuous on $[a, b]$. Therefore, the Mean Value theorem for derivatives (Theorem 3.38) states there is some $c$ such that

$$F'(c) = \frac{1}{b-a}\big(F(b) - F(a)\big). \tag{6.37}$$

Since $F$ is an anti-derivative of $f$ and $\int_a^a f(t)dt = 0$, we have

$$f(c) = F'(c) = \frac{1}{b-a}\left(\int_a^b f(t)dt - \int_a^a f(t)dt\right) = \frac{1}{b-a} \int_a^b f(x)dx. \tag{6.38}$$

$\square$

**Remark 6.39.** *The Mean Value Theorem for integrals gives a way of finding the average value of a function over an interval. The integral of $f$ amounts to summing up all the values $f$ on the interval $[a, b]$, then the factor of $\frac{1}{b-a}$ accounts for length of the interval that we summed over.*

Now that we've a solid grounding in how integration works and some of its important tools, we'll move to the kinds of problems in which integration can be applied.