

# Chapter 1

## Limits of Functions

At its heart, calculus is just the mathematics of change. In particular, calculus provides the tools necessary to describe how a function changes. Since a function can be used to represent many quantities that appear in real life, calculus therefore gives us a way to study how many different quantities (such as temperature, acceleration, energy of a signal, etc.) can change. But before we can start using calculus, we need to come up with the proper language to describe it. The language we will need to develop is that of the *limit of a function*.

### 1.1 The Limit of a Function

#### 1.1.1 Motivating Limits

Imagine creating a mathematical valley out of the graph of  $f(x) = x^2$ , and in this valley walks a mathematical ant. The ant is walking towards the place on the hill directly above  $x = -\frac{1}{2}$ . At  $x = -2$ , the ant is 4 units above the ground. At  $x = -1$ , the ant is now 1 unit above the ground. As the ant moves towards  $x = -\frac{1}{2}$ , its height above the ground gets closer and closer to  $\frac{1}{4}$ . To say it more clearly, and because modern calculus is built on this idea, *as the ant approaches* the point above  $x = -\frac{1}{2}$ , its *height above the ground approaches* the value  $\frac{1}{4}$ . In other words, the *limit* of  $f(x) = x^2$  as  $x$  approaches  $-\frac{1}{2}$  is  $\frac{1}{4}$ .

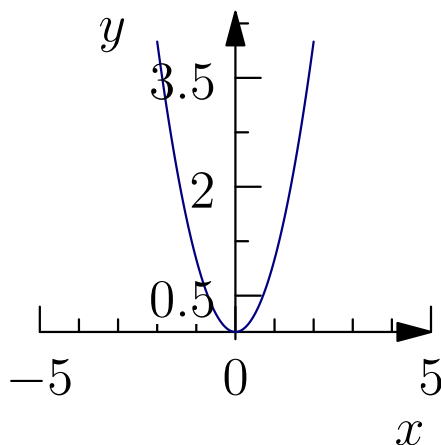


Figure 1.1.1: A mathematical valley minus a mathematical ant.

Remember that a function  $f$  of the variable  $x$  is just a rule that turns one number,  $x$ , into another number,  $f(x)$ . So the idea that the *limit of a function* is trying to express is what happens to the number  $f(x)$  (the output) as the number  $x$  (the input) approaches some particular value.

Actually, functions are much more general than this. But for calculus, it won't hurt us to view functions in this way.

We're not quite ready to define the limit of a function precisely, but we can point one thing out right away: *the limit of a function requires two pieces of information: the function itself and the number that  $x$  is approaching*. The limit should then be whatever number that  $f(x)$  is approaching.

**Example 1.1.2 Estimating the limit of a trigonometric function.** Let  $f(x) = \sin x \cos x$ . What is the limit of  $f(x)$  as  $x$  approaches the number  $\pi$ ?

**Solution.** We don't have a lot of tools to find limits yet, so we'll try to estimate it instead. What we'll do is we'll plug numbers that are closer and closer to  $\pi$  into  $f(x)$ . Let's list several values of  $f(x)$  as  $x$  gets closer to  $\pi$  from the left:

$x$	$f(x)$
3	-.140
3.1	-.042
3.14	-.002

We can even let  $x$  approach  $\pi$  from the other direction as well (i.e. "from the right") and  $f(x)$  will still approach 0 as  $x$  gets closer and closer to  $\pi$ . So it looks like the limit should be 0.

To keep ourselves from writing "the limit of  $f(x)$  as  $x$  approaches some number  $a$ " over and over, let's introduce some notation:  $\lim_{x \rightarrow a} f(x)$ .

**Example 1.1.3 Limit of a piecewise function.** Let

$$f(x) = \begin{cases} \sin x & \text{if } x < 0 \\ \cos x & \text{otherwise} \end{cases}.$$

Find  $\lim_{x \rightarrow 0} f(x)$ .

**Solution.** If we graph this function, we see that at  $x = 0$  there is a jump in the graph. In particular, if  $x$  approaches 0 from the left then  $f(x)$  approaches 0, whereas if  $x$  approaches 0 from the right then  $f(x)$  approaches 1. So this function does not appear to have an unambiguous limit as  $x$  approaches 0. Another way to say this:  $\lim_{x \rightarrow 0} f(x)$  *does not exist*.

[Example 1.1.3](#) shows us something very important about limits: they depend on the two different ways  $x$  can approach a number. So we introduce two new pieces of notation: the **left-hand limit**  $\lim_{x \rightarrow a^-} f(x)$  will stand for the value  $f(x)$  approaches (if any) as  $x$  approaches  $a$  from the left (i.e. as  $x$  increases to  $a$ ), and the **right-hand limit**  $\lim_{x \rightarrow a^+} f(x)$  will stand for the value  $f(x)$  approaches (if any) as  $x$  approaches  $a$  from the right (i.e. as  $x$  decreases to  $a$ ).

At this point, we can make a rough definition for the limit of a function.

**Definition 1.1.4 Limit of a Function.** Let  $f(x)$  be a function. Suppose that both the left-hand limit  $\lim_{x \rightarrow a^-} f(x)$  and the right-hand limit  $\lim_{x \rightarrow a^+} f(x)$

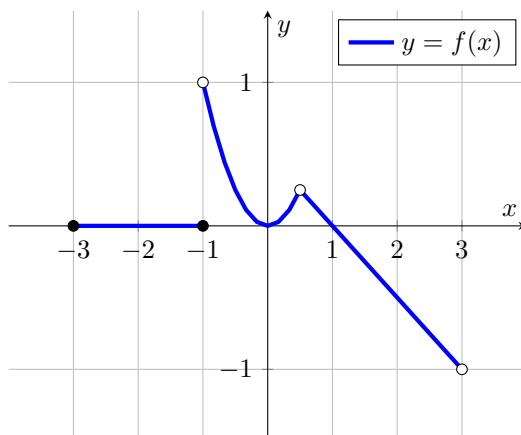
exist and are equal to the same number  $L$ . Then we say that the limit of  $f(x)$  as  $x$  approaches  $a$  exists and is equal to  $L$ . We denote this by writing  $\lim_{x \rightarrow a} f(x) = L$ .

**Example 1.1.5 Piecewise function again.** Let  $f(x)$  be given by

$$f(x) = \begin{cases} 0 & -3 \leq x \leq -1 \\ x^2 & -1 < x < \frac{1}{2} \\ \frac{1-x}{2} & \frac{1}{2} < x < 3. \end{cases}$$

Evaluate  $\lim_{x \rightarrow -1} f(x)$  and  $\lim_{x \rightarrow \frac{1}{2}} f(x)$ .

**Solution.** If we graph  $f(x)$ , we get the following:



The graph shows us that  $\lim_{x \rightarrow -1^-} f(x) = 0$ , while  $\lim_{x \rightarrow -1^+} f(x) = 1$ . Therefore  $\lim_{x \rightarrow -1} f(x)$  does not exist. On the other hand,  $\lim_{x \rightarrow \frac{1}{2}} f(x)$  exists and is equal to  $\frac{1}{4}$ .

## 1.2 Computing Limits

We've got a handle on how to estimate limits from [Section 1.1](#), but the process is very tedious. It requires either graphing the function in question or laboriously entering values into a calculator. So our first order of business now that we have the concept of a limit is to find an easier way to calculate it. This will be a running theme throughout the course.

### 1.2.1 The Limit Laws

In many cases of interest, we can use knowledge of simpler limits to obtain more complicated limits. We do this via the **Limit Laws**. Before we get to them, we'll state two very simple (and hopefully very believable) limits.

**Proposition 1.2.1 Simple Limits.** *For any value of  $a$ , the following limits hold:*

$$\lim_{x \rightarrow a} c = c$$

if  $c$  is a constant and

$$\lim_{x \rightarrow a} x = a.$$

**Theorem 1.2.2 The Limit Laws.** Let  $c$  be a constant, let  $n$  be a positive whole number and let  $f(x)$  and  $g(x)$  be functions. Suppose that  $\lim_{x \rightarrow a} f(x)$  and  $\lim_{x \rightarrow a} g(x)$  both exist for some number  $a$ . Then the following rules hold:

1. $\lim_{x \rightarrow a} [f(x) \pm g(x)] = \lim_{x \rightarrow a} f(x) \pm \lim_{x \rightarrow a} g(x)$	2. $\lim_{x \rightarrow a} cf(x) = c \lim_{x \rightarrow a} f(x)$
3. $\lim_{x \rightarrow a} [f(x)g(x)] = \lim_{x \rightarrow a} f(x) \lim_{x \rightarrow a} g(x)$	4. $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \frac{\lim_{x \rightarrow a} f(x)}{\lim_{x \rightarrow a} g(x)}$ (if $\lim_{x \rightarrow a} g(x) \neq 0$ )
5. $\lim_{x \rightarrow a} [f(x)^n] = [\lim_{x \rightarrow a} f(x)]^n$	6. $\lim_{x \rightarrow a} \sqrt[n]{f(x)} = \sqrt[n]{\lim_{x \rightarrow a} f(x)}$

**Table 1.2.3:** The Limit Laws.

Note that item six in the table above only holds (in this class...) if  $n$  is odd or if  $\lim_{x \rightarrow a} f(x) \geq 0$ .

**Theorem 1.2.2** gives us the ability to compute a wide variety of limits.

**Example 1.2.4 Limit of a rational function.** Let

$$f(x) = \frac{3 - x^5 + 5x}{2x - \sqrt[4]{x}}.$$

Evaluate  $\lim_{x \rightarrow 16} f(x)$ .

**Solution.** We can evaluate  $\lim_{x \rightarrow 16} f(x)$  by making use of the appropriate Limit Laws and **Proposition 1.2.1**:

$$\begin{aligned}
 \lim_{x \rightarrow 16} \frac{3 - x^5 + 5x}{2x - \sqrt[4]{x}} &= \frac{\lim_{x \rightarrow 16} (3 - x^5 + 5x)}{\lim_{x \rightarrow 16} (2x - \sqrt[4]{x})} && \text{by Limit Law 4} \\
 &= \frac{\lim_{x \rightarrow 16} 3 - \lim_{x \rightarrow 16} x^5 + \lim_{x \rightarrow 16} 5x}{\lim_{x \rightarrow 16} 2x - \lim_{x \rightarrow 16} \sqrt[4]{x}} && \text{by Limit Law 1.} \\
 &= \frac{3 - [\lim_{x \rightarrow 16} x]^5 + 5 \lim_{x \rightarrow 16} x}{2 \lim_{x \rightarrow 16} x - \sqrt[4]{\lim_{x \rightarrow 16} x}} && \text{by Limit Laws 2, 5, and 6.} \\
 &= \frac{3 - 16^5 + 80}{30}
 \end{aligned}$$

In particular, Limit Laws 1-5 give us the following: if  $f(x)$  is a polynomial or rational function, then  $\lim_{x \rightarrow a} f(x) = f(a)$  as long as  $a$  is in the domain of  $f(x)$ . If  $a$  is not in the domain, trickery may be required.

**Example 1.2.5 Trickery.** Evaluate

$$\lim_{x \rightarrow 1} \frac{\sqrt{x} - 1}{x - 1}.$$

**Solution.** First, note that we can't use the Limit Laws right away since the denominator is 0 at  $x = 1$ . What we need to do is use algebra to simplify the expression inside the limit:

$$\begin{aligned}
 \frac{\sqrt{x} - 1}{x - 1} &= \frac{\sqrt{x} - 1}{x - 1} \cdot \frac{\sqrt{x} + 1}{\sqrt{x} + 1} \\
 &= \frac{x - 1}{(x - 1)(\sqrt{x} + 1)} \\
 &= \frac{1}{\sqrt{x} + 1}.
 \end{aligned}$$

Now we can use the Limit Laws to find the limit as  $x$  approaches 1, since we no longer have a divide-by-zero problem in the denominator:

$$\lim_{x \rightarrow 1} \frac{\sqrt{x} - 1}{x - 1} = \lim_{x \rightarrow 1} \frac{1}{\sqrt{x} + 1}$$

$$= \frac{1}{2}.$$

## 1.3 Continuity

We saw in [Section 1.2](#) that for a function like  $f(x) = x + 3x^2$ , we could evaluate  $\lim_{x \rightarrow a} f(x)$  by simply plugging in  $x = a$ . In other words,  $\lim_{x \rightarrow a} f(x) = f(a)$ . Functions that have this property are extremely important in mathematics, so we give them a name.

**Definition 1.3.1 Continuous Functions.** Let  $f(x)$  be a function and suppose that  $a$  is in the domain of  $f(x)$ . Then we say that  $f(x)$  is **continuous at  $a$**  if

$$\lim_{x \rightarrow a} f(x) = f(a).$$

Otherwise, we say that  $f(x)$  is **discontinuous at  $a$** . We say that  $f(x)$  is continuous on an interval if it is continuous at every point of an interval. Otherwise, we say that  $f(x)$  is discontinuous on the interval.

[Definition 1.3.1](#) says that it is extremely easy to evaluate limits of continuous functions: just plug the value that  $x$  is approaching into the function  $f(x)$ . So the limit is then  $f(a)$ . If a function  $f(x)$  is continuous on an interval, then this means that the graph of  $f(x)$  has no "gaps" over this interval.

**Example 1.3.2 Determining if a function is continuous.** Let  $f(x) = \frac{1}{x}$ . Is  $f(x)$  continuous on  $(-\infty, \infty)$ ?

**Solution.** If we graph  $f(x)$ , then we see that it is discontinuous at  $x = 0$ . Therefore  $f(x)$  is discontinuous on the interval  $(-\infty, \infty)$ .

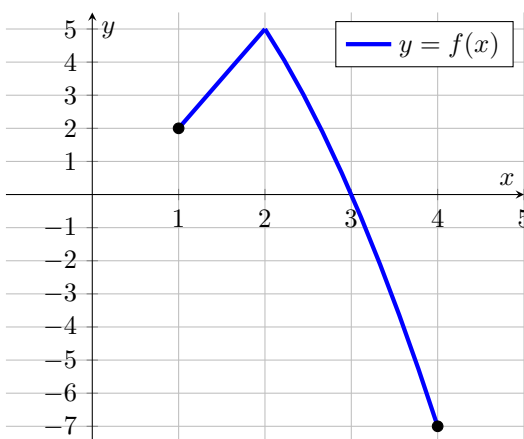
If we're dealing with a function on a (bounded) closed interval, we need to introduce some new terminology. We say that a function  $f(x)$  is **continuous from the left** at  $x = a$  if  $\lim_{x \rightarrow a^-} f(x) = f(a)$ . Similarly, we say that  $f(x)$  is **continuous from the right** at  $x = a$  if  $\lim_{x \rightarrow a^+} f(x) = f(a)$ . This is of course assuming that  $a$  is in the domain of  $f(x)$ . Finally, we say that  $f(x)$  is continuous on the closed interval  $[a, b]$  if it is continuous on  $(a, b)$ , continuous from the right at  $a$  and continuous from the left at  $b$ . The main idea is still that the graph contains no gaps over this interval.

**Example 1.3.3 Continuity over a closed interval.** Let  $f(x)$  be given by

$$f(x) = \begin{cases} 3x - 1 & 1 \leq x < 2 \\ 9 - x^2 & 2 \leq x \leq 4. \end{cases}$$

Is  $f(x)$  continuous over  $[1, 4]$ ?

**Solution.** If we graph  $f(x)$  over this interval, we get the following:



So from the graph it appears that  $f(x)$  is continuous on this interval.

Remember that we said a function is continuous over an interval if its graph has no gaps over that interval. This is a very rough explanation of continuity, but it should make the following theorem plausible.

**Theorem 1.3.4 Continuous Functions.** *Polynomial, rational, root and trigonometric functions are continuous at every point of their domain.*

Although it doesn't directly mention piecewise functions, [Theorem 1.3.4](#) is still useful for determining if they are continuous. If a piecewise function is defined using any of the functions from [Theorem 1.3.4](#), then the only points we really need to check for continuity are the the places where the function "changes rules".

**Example 1.3.5 Another piecewise function.** Over what intervals is the function  $g(s)$  given by

$$g(s) = \begin{cases} s^2 & s < -1 \\ -5 \cos(\pi s) & -1 \leq s \leq 1 \\ 5 & s > 1 \end{cases}$$

continuous?

**Solution.** We need to check continuity at  $s = -1$  and  $s = 1$ . At  $s = -1$ , we need to make sure that  $\lim_{s \rightarrow -1} g(s)$  exists and is equal to  $g(-1)$ . Since

$$\lim_{s \rightarrow -1^-} g(s) = \lim_{s \rightarrow -1^-} s^2 = 1$$

and

$$\lim_{s \rightarrow -1^+} g(s) = \lim_{s \rightarrow -1^+} \cos(\pi s) = -1,$$

it follows that  $\lim_{s \rightarrow -1} g(s)$  does not exist. So  $g(s)$  can't be continuous at  $s = -1$ .

On the other hand, since  $\lim_{s \rightarrow 1^-} g(s) = 5 = \lim_{s \rightarrow 1^+} g(s)$ , it follows that  $\lim_{s \rightarrow 1} g(s)$  exists and is equal to 5. Since  $g(1)$  also equals 5,  $g(s)$  is continuous at  $s = 1$ . So  $g(s)$  must be continuous on  $(-\infty, -1) \cup (-1, \infty)$ .

We can also build more complicated continuous functions out of simpler ones.

**Theorem 1.3.6 Combining Continuous Functions.** *Let  $f(x)$  and  $g(x)$  be continuous at a point  $a$ . Then the following statements are true:*

1.  $f(x) + g(x)$  is continuous at  $a$ .
2.  $f(x)g(x)$  is continuous at  $a$ .
3.  $\frac{f(x)}{g(x)}$  is continuous at  $a$  if  $g(a) \neq 0$ .
4.  $f(g(x))$  is continuous at  $a$  if  $g(a)$  is in the domain of  $f(x)$ .

**Example 1.3.7 Determining where functions are continuous.** Let  $h(t) = \sqrt{t} + \frac{t}{t-4} - \frac{t+1}{t^2-1}$ . On what intervals is  $h(t)$  continuous?

**Solution.** By [Theorem 1.3.6](#),  $h(t)$  should be continuous wherever  $\sqrt{t}$ ,  $\frac{t}{t-4}$  and  $\frac{t+1}{t^2-1}$  are all defined. Since  $\sqrt{t}$  is defined for  $t \geq 0$ ,  $\frac{t}{t-4}$  is defined for  $t \neq 4$  and  $\frac{t+1}{t^2-1}$  is defined for  $t \neq \pm 1$ , it follows that  $h(t)$  is continuous on  $[0, 1) \cup (1, 4) \cup (4, \infty)$ .

**Example 1.3.8 Using continuity to evaluate a limit.** Evaluate  $\lim_{x \rightarrow \pi} \cos(x + \sin x)$ .

**Solution.** Since  $x$ ,  $\cos x$  and  $\sin x$  are all continuous, this means that  $\cos(x + \sin x)$  must be continuous as well. Therefore

$$\lim_{x \rightarrow \pi} \cos(x + \sin x) = \cos(\pi + \sin \pi) = -1.$$

## 1.4 Limits Involving Infinity

### 1.4.1 Limits Involving Vertical Asymptotes

Consider the function  $f(x) = \frac{1}{x}$ . We know from algebra that this function has a vertical asymptote at  $x = 0$ , and so in particular is undefined there. However, just because it's undefined at  $x = 0$  doesn't mean that we can't gather important information about the function near 0. This is because the function behaves in a very specific way as we let  $x$  approach 0. For example, if we let  $x$  approach 0 from the right, then  $f(x)$  increases without bound. Similarly,  $f(x)$  decreases without bound as  $x$  approaches 0 from the left.

Even though  $f(x)$  is not approaching a specific value as  $x$  approaches 0 from either direction, this behavior shows up often enough and is important enough that we want to introduce notation to describe it. For this function, we would say  $\lim_{x \rightarrow 0^+} \frac{1}{x} = \infty$  and  $\lim_{x \rightarrow 0^-} \frac{1}{x} = -\infty$ .

Now consider  $g(x) = \frac{1}{(x-3)^2}$ . Then  $\lim_{x \rightarrow 3^-} g(x) = \lim_{x \rightarrow 3^+} g(x) = \infty$  since the function increases without bound when  $x$  approaches 3 from both directions. In this case, we say that  $\lim_{x \rightarrow 3} g(x) = \infty$ .

It's *extremely* important to remember that the symbol  $\infty$  is not being used to represent a number or variable that we can perform algebra on. It's a symbol indicating how a particular function is behaving at a certain point.

If  $f(x)$  is a function and  $\lim_{x \rightarrow a} f(x) = \pm\infty$ ,  $\lim_{x \rightarrow a^-} f(x) = \pm\infty$  or  $\lim_{x \rightarrow a^+} f(x) = \pm\infty$ , then this means that the function has a vertical asymptote at  $x = a$ . In this course, this basically corresponds to a divide-by-zero problem.

**Example 1.4.1 Infinite limit involving a rational function.** Determine  $\lim_{x \rightarrow 4^-} \frac{x-3}{2x-8}$ .

**Solution.** If we try to plug in  $x = 4$  into  $\frac{x-3}{2x-8}$  we get  $\frac{1}{0}$ , which means we have run into a divide-by-zero problem. This is a good hint that the limit

should be  $\pm\infty$ , we just need to figure out the correct sign. There are a couple ways we can do this. First, we could set up a sign chart for this function to see where it's positive and negative and then use that to see if it's increasing or decreasing without bound as  $x \rightarrow 4^-$ . Second, we could just plug in values of  $x$  that are closer and closer to 4 and see how the function behaves. Either way, we see that it's negative for values of  $x$  that are close to (but less than) 4. Hence  $\lim_{x \rightarrow 4^-} \frac{x-3}{2x-8} = -\infty$ .

### 1.4.2 Limits at Infinity

[Subsection 1.4.1](#) involved limits of functions whose values approached  $\pm\infty$ . Now we look at what can happen to a function if its input values approach  $\pm\infty$ . First, a definition of sorts.

**Definition 1.4.2 Limit at Infinity.** Let  $f(x)$  be a function. We say that  $\lim_{x \rightarrow \infty} f(x) = L$  if  $f(x)$  gets (and stays) arbitrarily close to  $L$  as  $x$  is made arbitrarily large. Similarly, we say that  $\lim_{x \rightarrow -\infty} f(x) = L$  if  $f(x)$  gets (and stays) arbitrarily close to  $L$  as  $x$  is made arbitrarily small.

**Example 1.4.3 An important limit at infinity.** Let  $f(x) = \frac{1}{x}$ . Determine  $\lim_{x \rightarrow \infty} f(x)$ .

**Solution.** As  $x$  gets arbitrarily large,  $\frac{1}{x}$  gets arbitrarily close to 0. Therefore  $\lim_{x \rightarrow \infty} f(x) = 0$ .

[Example 1.4.3](#) holds for many other reciprocal powers of  $x$ . In particular, if  $n > 0$  then  $\lim_{x \rightarrow \infty} \frac{1}{x^n} = \lim_{x \rightarrow -\infty} \frac{1}{x^n} = 0$ .

**Example 1.4.4 A limit at infinity involving cosine.** Let  $g(s) = \cos s^{-3}$ . Compute  $\lim_{s \rightarrow -\infty} g(s)$ .

**Solution.** First, note that  $g(s) = \cos \frac{1}{s^3}$ . By the previous remark, we know that  $\lim_{s \rightarrow -\infty} \frac{1}{s^3} = 0$ . Therefore  $\lim_{s \rightarrow -\infty} g(s) = \cos 0 = 1$ .

The reason we were able to find the limit in [Example 1.4.4](#) was because of the following fact: if  $\lim_{x \rightarrow a} g(x)$  exists and if  $f(x)$  is continuous at  $\lim_{x \rightarrow a} g(x)$ , then  $\lim_{x \rightarrow a} f(g(x)) = f(\lim_{x \rightarrow a} g(x))$ . Basically, we can swap continuous functions with limits without causing any harm.

For limits at infinity involving powers of a variable, it is the highest power variables that determine the outcome.

**Example 1.4.5 Limit at infinity of a rational function.** Let

$$f(t) = \frac{100t - 3 + t^3}{5t^2 + 1} \text{ and } g(t) = \frac{1 - 2t^4}{5t^4 + 3000t}.$$

Find  $\lim_{t \rightarrow \infty} f(t)$  and  $\lim_{t \rightarrow -\infty} g(t)$ .

**Solution.** Let's start with  $f(t)$ . To figure out what this limit should be, we could try the following. As  $t$  gets very large the  $t^3$  term in the numerator should drown out everything else in the numerator. Similarly, the  $5t^2$  term in the denominator should drown out everything else in the denominator. So for  $t$  very large,  $f(t) \approx \frac{t^3}{5t^2} = \frac{1}{5}t$ . Hence  $f(t)$  should probably go to  $\infty$  as  $t$  goes to  $\infty$ . To make this more precise, we'll just divide the numerator and denominator by the largest power of the denominator,  $t^2$ , and then take the limit:

$$\lim_{t \rightarrow \infty} f(t) = \lim_{t \rightarrow \infty} \frac{100t - 3 + t^3}{5t^2 + 1}$$



$$\begin{aligned}
&= \lim_{t \rightarrow \infty} \frac{\frac{100}{t} - \frac{3}{t^2} + t}{5 + \frac{1}{t^2}} \\
&= \infty.
\end{aligned}$$

We can find  $\lim_{t \rightarrow -\infty} g(t)$  using the same idea. Just divide by the highest power in the denominator and then take the limit:

$$\begin{aligned}
\lim_{t \rightarrow -\infty} g(t) &= \lim_{t \rightarrow -\infty} \frac{\frac{1}{t^4} - 2}{5 + \frac{3000}{t^3}} \\
&= -\frac{2}{5}.
\end{aligned}$$

These limits at infinity have a graphical meaning as well. If  $\lim_{x \rightarrow \infty} f(x)$  or  $\lim_{x \rightarrow -\infty} f(x)$  exists and is equal to  $L$ , then the line  $y = L$  is a horizontal asymptote of the graph of  $y = f(x)$ .

**Example 1.4.6 Asymptotic equivalence.** Two functions  $f(x)$  and  $g(x)$  are said to be **asymptotically equivalent**, written  $f \sim g$ , if the following is true:

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 1.$$

Show that  $\sin \frac{1}{x} \sim \frac{1}{x}$ .

**Solution.** All we need to do is compute  $\lim_{x \rightarrow \infty} \frac{\sin \frac{1}{x}}{\frac{1}{x}}$ :

$$\begin{aligned}
\lim_{x \rightarrow \infty} \frac{\sin \frac{1}{x}}{\frac{1}{x}} &= \lim_{\theta \rightarrow 0^+} \frac{\sin \theta}{\theta} \\
&= 1.
\end{aligned}$$

Therefore  $\sin \frac{1}{x} \sim \frac{1}{x}$ .



# Chapter 2

## Derivatives

There are two problems that inspired the creation of calculus. The first is the tangent line problem, or rate of change problem. This problem is concerned with determining how a quantity  $f(x)$  changes as  $x$  varies. We will focus on this problem and its solution, derivatives, for the next two chapters. Afterwards, we will focus on the area problem, the second problem that inspired the creation of calculus.

### 2.1 The Definition of the Derivative

#### 2.1.1 Tangent Lines

Consider  $f(x) = x^2$ . If we graph this, we get a parabola. What we'd like to do is to find a way to describe how quickly this parabola is changing at a point, i.e. find the "slope" of the parabola. One way to try to deal with this is to use **secant lines**. Recall that the secant line through the points  $(a, f(a))$  and  $(b, f(b))$  has slope  $\frac{f(b)-f(a)}{b-a}$ , which is just the average rate of change of  $f$  from  $x = a$  to  $x = b$ . If  $b$  is very close to  $a$ , then the slope of the secant line through these points should be a good approximation to the "slope" of  $f(x) = x^2$  at the point  $x = a$ .

**Example 2.1.1 Secant lines on a parabola.** Let  $f(x) = x^2$ . Find the slope of the secant line through  $(2, f(2))$  and  $(x, f(x))$ .

**Solution.** Since the slope of the secant line is the average rate of change, we get that the slope must be equal to

$$\frac{f(x) - f(2)}{x - 2} = \frac{x^2 - 4}{x - 2} = x + 2.$$

What [Example 2.1.1](#) is telling us is that if  $x \approx 2$ , then the slope of  $f(x) = x^2$  at  $x = 2$  should be very close to  $x + 2$ , the slope of the secant line. Now we'll do a trick that shows up everywhere in calculus, and is the entire reason we introduced limits in the first place. We'll make this approximation exact by taking a limit. In particular, we'll say that the slope of  $f(x) = x^2$  should be equal to

$$\lim_{x \rightarrow 2} \frac{f(x) - f(2)}{x - 2} = \lim_{x \rightarrow 2} (x + 2) = 4.$$

This is the slope of the **tangent line** to  $f(x) = x^2$  at  $x = 2$ . Instead of an average rate of change over an interval  $[2, x]$ , we now have an **instantaneous rate of change** at a point  $x = 2$ .

**Definition 2.1.2 Tangent Lines.** The tangent line to a curve  $y = f(x)$  through a point  $(a, f(a))$  is the line passing through  $(a, f(a))$  with slope given by

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a},$$

assuming this limit exists. The slope of the tangent line represents the slope of the graph of  $f(x)$  at  $a$  and gives the instantaneous rate of change of  $f(x)$  at  $x = a$ .

**Example 2.1.3 Tangent line to a root.** Find the equation of the tangent line to  $y = \sqrt{x}$  through the point  $(4, 2)$ .

**Solution.** We need two things to find the equation of a line: the slope of the line and a point on the line. Since we know the tangent line has to pass through  $(4, 2)$ , we just need to find the slope. The slope is given by

$$\lim_{x \rightarrow 4} \frac{\sqrt{x} - 2}{x - 4} = \lim_{x \rightarrow 4} \frac{x - 4}{(x - 4)(\sqrt{x} + 2)} = \frac{1}{4}.$$

Hence the equation of the tangent line through  $(4, 2)$  is

$$y - 2 = \frac{1}{4}(x - 4).$$

As a reminder, the slope of the tangent line represents the slope, or instantaneous rate of change, of the function.

**Example 2.1.4 Velocity from position.** The displacement (i.e. position) of a particle moving in a straight line is described by the function  $s(t) = 3t^2 - 5t$ , where  $t$  is in seconds and  $s$  is in meters. Find the velocity, or instantaneous rate of change, of the particle at  $t = 3$ .

**Solution.** The velocity is just the slope of the tangent line to  $s(t)$  at  $t = 3$ , which we can find as follows:

$$\begin{aligned} \lim_{t \rightarrow 3} \frac{s(t) - s(3)}{t - 3} &= \lim_{t \rightarrow 3} \frac{(3t^2 - 5t) - 12}{t - 3} \\ &= \lim_{t \rightarrow 3} \frac{3t^2 - 5t - 12}{t - 3} \end{aligned}$$

At this step it's a little unclear where to go next, so we'll try long division. If we do so, we get

$$\frac{3t^2 - 5t - 12}{t - 3} = 3t + 4.$$

Hence the velocity must be

$$\lim_{t \rightarrow 3} (3t + 4) = 13$$

meters per second.

[Example 2.1.4](#) was a little tricky because we needed to compute  $\lim_{t \rightarrow 3} \frac{3t^2 - 5t - 12}{t - 3}$ , and it was unclear how to simplify this at first. This stemmed in large part from how we computed the velocity in the first place, using the formula

$$\lim_{t \rightarrow 3} \frac{s(t) - s(3)}{t - 3}$$

or more generally

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}.$$

We want to rewrite this formula to make it a little easier to work with in certain cases. We'll do this by making the denominator easier to handle. In particular, set  $x - a = h$ . Then

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

Either formula can be used to compute the slope of the tangent line.

**Example 2.1.5 Velocity revisited.** Let  $s(t)$  be given as in [Example 2.1.4](#). Find the velocity at  $t = 3$ .

**Solution.** We know that the velocity should be 13, but we'll try to find it again using our new formula. So the velocity should also be

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{s(3 + h) - s(3)}{h} &= \lim_{h \rightarrow 0} \frac{3(9 + 6h + h^2) - 5(3 + h) - 12}{h} \\ &= \lim_{h \rightarrow 0} \frac{13h + 3h^2}{h} \\ &= \lim_{h \rightarrow 0} (13 + 3h) \\ &= 13. \end{aligned}$$

Typically, if  $f(x) - f(a)$  is easy to factor in terms of  $x - a$  we'll want to use the first formula we had for computing rates of change. Otherwise, we'll stick with the new formula involving  $h$ .

### 2.1.2 The Derivative

Suppose we go back to [Example 2.1.4](#) one more time, but now we want to find the velocity of  $s(t)$  at an arbitrary number  $a$ . Then we could still use our limit formulas, which would give us

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{s(a + h) - s(a)}{h} &= \lim_{h \rightarrow 0} \frac{3(a^2 + 2ah + h^2) - 5(a + h) - 3a^2 + 5a}{h} \\ &= \lim_{h \rightarrow 0} \frac{6ah + 3h^2 - 5h}{h} \\ &= \lim_{h \rightarrow 0} (6a + 3h - 5) \\ &= 6a - 5. \end{aligned}$$

So the velocity at  $t = a$  of the particle modeled by  $s(t)$  is given by  $6a - 5$ . So we can represent the velocity, or rate of change or slope of the tangent line, by a function. We call this the **derivative**.

**Definition 2.1.6 Definition of the Derivative.** Let  $f(x)$  be a function. Then its derivative at  $x = a$  is the number  $f'(a)$  given by

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

assuming the limit exists. If this limit exists, we say that the function  $f(x)$  is **differentiable** at  $a$ .

We could also define the derivative by

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}.$$

These two limits are equivalent.

The derivative of a function  $f(x)$  at a point  $a$  represents two things: the slope of the tangent line to  $f(x)$  at  $a$  and the instantaneous rate of change of  $f(x)$  at  $a$ .

**Example 2.1.7 Slope of the sine function.** Let  $f(x) = \sin x$ . Find its slope at 0.

**Solution.** The slope at 0 is exactly  $f'(0)$ , which is

$$\lim_{h \rightarrow 0} \frac{\sin(0 + h) - \sin(0)}{h} = \lim_{h \rightarrow 0} \frac{\sin h}{h} = 1.$$

**Example 2.1.8 Tangent line of the sine function.** Find the equation of the tangent line to  $f(x) = \sin x$  at 0.

**Solution.** The tangent line must pass through  $(0, 0)$  and must have slope  $f'(0) = 1$ , so its equation is

$$y - 0 = 1(x - 0)$$

or just  $y = x$ .

## 2.2 The Derivative as a Function

### 2.2.1 The Derivative Function

There's no reason we can't look at an arbitrary value for  $a$  in the definition of  $f'(a)$  given in [Definition 2.1.6](#). If we do this, we can define the *derivative function*.

**Definition 2.2.1 The Derivative Function.** Let  $f(x)$  be a function. The **derivative function**, or more simply **derivative**, of  $f(x)$  is the function  $f'(x)$  defined by

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

assuming this limit exists. This is also often denoted by  $\frac{d}{dx}(f)$  or  $\frac{df}{dx}$ . If this limit exists for all  $x$  in some interval  $I$ , we say that  $f$  is **differentiable on  $I$** , or more simply **differentiable** if we do not wish to specify the interval.

**Example 2.2.2 Computing a derivative.** Compute the derivative of  $f(x) = x - 3x^2$ .

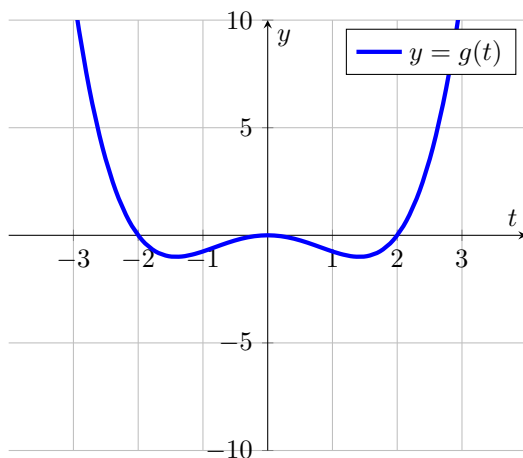
**Solution.** Using [Definition 2.2.1](#), we have

$$\begin{aligned} f'(x) &= \lim_{h \rightarrow 0} \frac{[(x + h) - 3(x + h)^2] - [x - 3x^2]}{h} \\ &= \lim_{h \rightarrow 0} \frac{[x + h - 3x^2 - 6xh - 3h^2] - [x - 3x^2]}{h} \\ &= \lim_{h \rightarrow 0} \frac{h - 6xh - 3h^2}{h} \end{aligned}$$

$$\begin{aligned}
 &= \lim_{h \rightarrow 0} (1 - 6x - 3h) \\
 &= 1 - 6x.
 \end{aligned}$$

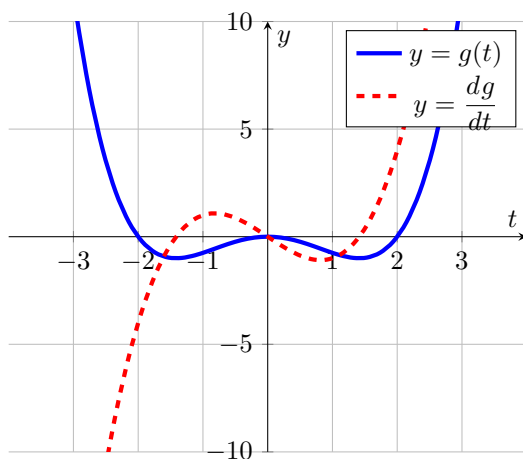
If  $f(x)$  is a function, then its derivative  $f'(x)$  (assuming it exists!) is a function that gives the rate of change of  $f$  at  $x$ , or equivalently the slope of the tangent line to  $f$  at  $x$ .

**Example 2.2.3 Sketching a derivative.** A function  $g(t)$  is given by the following graph:



Sketch  $\frac{dg}{dt}$ .

**Solution.** Remember that  $\frac{dg}{dt}$  represents the slope of  $g(t)$ , so sketching  $\frac{dg}{dt}$  amounts to sketching the different values that the slopes of  $g(t)$  can take. We can eyeball these values from Figure ?? . A rough sketch of  $\frac{dg}{dt}$ , added to the original graph, may look like the following:



We've mentioned before that continuous functions are functions whose graphs can be drawn without lifting your pencil off of the page. Likewise, differentiable functions are functions whose graphs can be drawn "smoothly",

without any sudden movements or cusps, and without drawing a vertical tangent line. If we think about these two concepts, we may suspect that a differentiable function is also continuous. If we can draw a graph smoothly, we certainly can't lift our pencil off the page to draw it. The next theorem makes this precise.

**Theorem 2.2.4 Differentiable Functions Are Continuous.** *Let  $f(x)$  be a function that is differentiable at  $x = a$ . Then  $f(x)$  is continuous at  $x = a$ .*

*Proof.* We need to show that  $\lim_{x \rightarrow a} f(x)$  exists and is equal to  $f(a)$ . To do this, we'll start by considering (somewhat counterintuitively)  $\lim_{x \rightarrow a} [f(x) - f(a)]$ :

$$\begin{aligned}\lim_{x \rightarrow a} [f(x) - f(a)] &= \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} (x - a) \\ &= \lim_{x \rightarrow a} f'(a)(x - a) \\ &= 0.\end{aligned}$$

Note that we are using our alternate definition of the derivative here.

Now we can prove that  $\lim_{x \rightarrow a} f(x) = f(a)$  as follows:

$$\begin{aligned}\lim_{x \rightarrow a} f(x) &= \lim_{x \rightarrow a} [f(x) - f(a) + f(a)] \\ &= \lim_{x \rightarrow a} [f(x) - f(a)] + \lim_{x \rightarrow a} f(a) \\ &= 0 + f(a).\end{aligned}$$

So  $\lim_{x \rightarrow a} f(x) = f(a)$ , which means that  $f(x)$  is continuous at  $a$ .  $\square$

At this point we might think that a continuous function should also be differentiable, but this is not the case.

**Example 2.2.5 A continuous function that is not differentiable at a point.** Let  $f(x) = |x|$ . Show that  $f$  is *not* differentiable at 0.

**Solution.** If we graph  $f(x)$  it looks like it shouldn't be differentiable at 0 because of the cusp. We'll try to prove this mathematically by showing that the limit in Definition 2.2.1 doesn't exist if  $x = 0$ . First, we'll compute the left hand limit:

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

Now, the right hand limit:

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

Since these limits are different,  $\lim_{h \rightarrow 0} \frac{f(0+h) - f(0)}{h}$  does not exist. Hence  $f$  is not differentiable at 0.

You may think that a continuous function must at least be differentiable "almost everywhere" at this point. After all, how could it be possible to draw a graph without lifting your pencil off the paper that still has a cusp or a vertical tangent line *everywhere*? Most mathematicians before the 19<sup>th</sup> century thought this as well, until Weierstrass came up with a function, the [Weierstrass function](#), that is continuous everywhere but differentiable *nowhere*.



### 2.2.2 Higher Order Derivatives

**Example 2.2.6 Acceleration from position.** The position of some particle moving in a line is given by  $s(t) = 3t - 5t^3$ , where  $t$  is in seconds and  $s$  is in meters. Find  $a(t)$ , the acceleration of the particle at time  $t$ .

**Solution.** Acceleration is the rate of change of velocity, and velocity is the rate of change of position. So we should probably find the velocity first! Let's call it  $v(t)$ . We have

$$\begin{aligned} v(t) &= \lim_{h \rightarrow 0} \frac{s(t+h) - s(t)}{h} \\ &= \lim_{h \rightarrow 0} \frac{3(t+h) - 5(t+h)^3 - 3t + 5t^3}{h} \\ &= \lim_{h \rightarrow 0} \frac{3h + 5[t^3 - (t+h)^3]}{h} \\ &= \lim_{h \rightarrow 0} \frac{3h + 5[(-h)(t^2 + t(t+h) + (t+h)^2)]}{h} \\ &= \lim_{h \rightarrow 0} (3 - 5[t^2 + t(t+h) + (t+h)^2]) \\ &= 3 - 15t^2 \end{aligned}$$

Now we can get the acceleration as well:

$$a(t) = \lim_{h \rightarrow 0} \frac{v(t+h) - v(t)}{h} = -30t.$$

In [Example 2.2.6](#), we had to take two derivatives of the original function  $s(t)$  in order to get the acceleration  $a(t)$ . In other words, *acceleration is the second derivative of position*. So  $a(t) = \frac{d}{dt} \frac{ds}{dt}$ , which we also write as  $\frac{d^2s}{dt^2}$  or  $s''(t)$ . This is an example of a **second-order derivative**. In general, we have the following definition.

**Definition 2.2.7  $n^{\text{th}}$ -order Derivatives.** Let  $f(x)$  be a function. The  $n^{\text{th}}$ -order derivative of  $f(x)$  is the function obtained by differentiating  $f(x)$   $n$  times. This function is denoted by

$$f^{(n)}(x) \quad \text{or} \quad \frac{d^n f}{dx^n}.$$

If  $n = 1, 2$  or  $3$ , we typically write  $f'(x)$ ,  $f''(x)$ , and  $f'''(x)$  instead of  $f^{(1)}(x)$ ,  $f^{(2)}(x)$  or  $f^{(3)}(x)$ .

Although it gets more difficult to assign a physical or geometric significance to higher order derivatives, we can still derive meaning from the second derivative. One interpretation of the second derivative is as acceleration, as shown in [Example 2.2.6](#), and it turns out there's a nice geometric interpretation as well. Recall that if  $f(x)$  is a function then  $f'(x)$  represents the slope, or rate of change, of the graph of  $f(x)$  at  $x$ . Therefore  $f''(x)$  represents the rate of change of the slope, i.e. how quickly the slope is increasing or decreasing. If  $f''(x) > 0$  then the slope of  $f(x)$  should be increasing, leading to a u-shaped graph. Conversely, if  $f''(x) < 0$  then the slope of  $f(x)$  should be decreasing, leading to an upside down u-shaped graph. This leads to the following definition.

**Definition 2.2.8 Concavity.** Let  $f(x)$  be a function with second derivative  $f''(x)$ . We say that  $f(x)$  is **concave up** (respectively, **concave down**) on an interval if  $f''(x) > 0$  (respectively,  $f''(x) < 0$ ) on that interval.

So functions that are concave up on an interval tend to be u-shaped on that interval, and functions that are concave down tend to be upside down u-shaped. See [Figure 2.2.8](#).

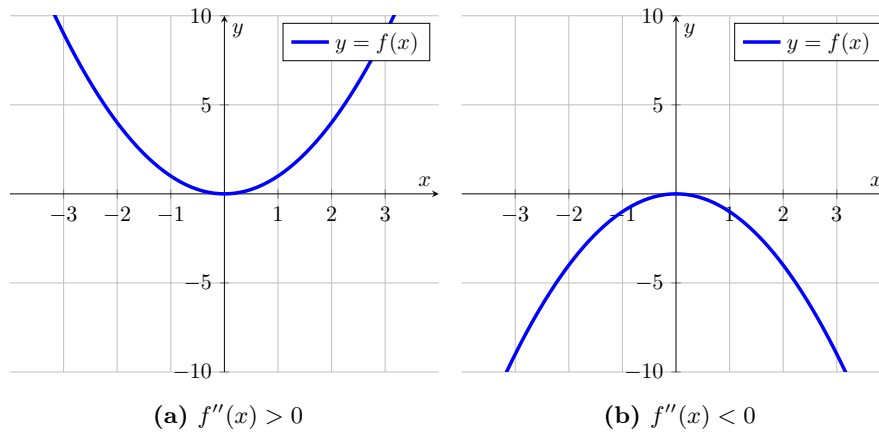


Figure 2.2.8: Concavity

## 2.3 Differentiation Formulas

Now we start to find methods that allow us to compute derivatives without going back to [Definition 2.2.1](#). Perhaps the easiest rule is the **constant rule**, which just says that the derivative of a constant is 0. We'll derive more complicated rules in this section and the next two.

### 2.3.1 The Power Rule and Trigonometric Derivatives

Our first goal will be to determine a general formula for the derivative of  $x^n$  for some power  $n$ . If  $n \geq 0$  is a whole number, then we can find the derivative of  $x^n$  without too much trouble. In fact, from the [Definition 2.2.1](#) it's not too hard to show that the derivative of  $x$  is 1, the derivative of  $x^2$  is  $2x$ , the derivative of  $x^3$  is  $3x^2$ , and so on. This suggests our first derivative rule, the **power rule**.

**Theorem 2.3.1 The Power Rule.** *Let  $f(x) = x^n$  where  $n$  is some real number. Then  $f'(x) = nx^{n-1}$ .*

Note that [Theorem 2.3.1](#) is actually quite general: it works for all powers of  $x$ ! This includes negative powers and fractional powers.

**Example 2.3.2 Derivatives using the power rule.** Find the derivatives of the following functions:

1.  $f(x) = x^7$ .
2.  $g(x) = \frac{1}{\sqrt[5]{x^{11}}}$ .
3.  $h(x) = x^\pi$ .

**Solution.** The derivative of  $f(x)$  isn't too hard to find using the power rule, and we quickly get  $f'(x) = 7x^6$ . For  $g(x)$ , first rewrite it as  $g(x) = x^{-\frac{11}{5}}$ . Then  $g'(x) = -\frac{11}{5}x^{-\frac{11}{5}-1}$ . Finally,  $h'(x) = \pi x^{\pi-1}$ .

With a little bit of geometry and the squeeze theorem, we can get the derivatives of the basic trigonometric functions  $\sin x$  and  $\cos x$ .

**Theorem 2.3.3 Derivatives of Sine and Cosine.** *Let  $x$  be in radians. Then*

$$\frac{d}{dx} \sin x = \cos x \quad \text{and} \quad \frac{d}{dx} \cos x = -\sin x.$$

Note that if  $x$  is in degrees instead of radians these formulas don't work. Instead, they become

$$\frac{d}{dx} \sin x = \frac{\pi}{180} \cos x \quad \text{and} \quad \frac{d}{dx} \cos x = -\frac{\pi}{180} \sin x.$$

**Example 2.3.4 Concavity of the sine function.** On which intervals is  $f(x) = \sin x$  concave up?

**Solution.** We need to find where  $f''(x)$  is positive. Since  $f''(x) = -\sin x$ , this means we need to figure out where  $\sin x$  is *negative*. If we go back to the unit circle definition of sine, then we can see that  $\sin x < 0$  on the following intervals:

$$\dots, (-3\pi, -2\pi), (-\pi, 0), (\pi, 2\pi), (3\pi, 4\pi), \dots$$

So  $\sin x$  is concave up on every open interval of the form  $((2k-1)\pi, 2k\pi)$  where  $k$  is some integer.

### 2.3.2 Derivatives of Sums and Constant Multiples

Now that we have derivative formulas for some basic functions, we want to extend these to more complicated functions. For this section we'll look at what happens when we multiply a function by a constant or add it to another function. In the next two sections we'll consider more advanced rules.

**Theorem 2.3.5 Constant Multiple Rule.** *Let  $f(x)$  be differentiable function and let  $c$  be some constant. Then  $\frac{d}{dx}[cf(x)] = cf'(x)$ .*

*Proof.* To prove this, we go back to [Definition 2.2.1](#):

$$\begin{aligned} \frac{d}{dx}[cf(x)] &= \lim_{h \rightarrow 0} \frac{cf(x+h) - cf(x)}{h} \\ &= c \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \\ &= cf'(x). \end{aligned}$$

Hence the derivative of  $cf(x)$  is  $cf'(x)$ . □

We can find the derivative of the sum of two functions just as easily.

**Theorem 2.3.6 Sum Rule.** *Let  $f(x)$  and  $g(x)$  be two differentiable functions. Then  $[f(x) + g(x)]' = f'(x) + g'(x)$ .*

## 2.4 Product and Quotient Rules

### 2.4.1 Product Rule

Many functions can be written in the form  $f(x)g(x)$ , where  $f(x), g(x)$  may each have previously known derivatives. What we want to do now is to find a way to get the derivative of  $f(x)g(x)$  from the derivatives of  $f(x)$  and  $g(x)$ . We do this using the **product rule**.

**Theorem 2.4.1 The Product Rule.** *Let  $f(x)$  and  $g(x)$  be differentiable functions. Then*

$$\frac{d}{dx}[f(x)g(x)] = f(x)g'(x) + f'(x)g(x).$$

*Proof.* We prove this using the definition of the derivative:

$$\begin{aligned} \frac{d}{dx}[f(x)g(x)] &= \lim_{h \rightarrow 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(x+h)g(x+h) - f(x+h)g(x) + f(x+h)g(x) - f(x)g(x)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(x+h)[g(x+h) - g(x)] + g(x)[f(x+h) - f(x)]}{h} \\ &= \lim_{h \rightarrow 0} f(x+h) \lim_{h \rightarrow 0} \frac{g(x+h) - g(x)}{h} + g(x) \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \\ &= f(x)g'(x) + g(x)f'(x). \quad \square \end{aligned}$$

**Example 2.4.2 Using the product rule.** Let  $J(v) = (v^{10} - 2 \sin v)(\cos v + \frac{1}{\sqrt[3]{x^4}})$ . Find  $J'(v)$ .

**Solution.** We could foil this out and take derivatives, but it will be easier to use the product rule.

### 2.4.2 The Quotient Rule

Now that we know how to differentiate products, we move on to quotients.

**Theorem 2.4.3 The Quotient Rule.** *Let  $f(x)$  and  $g(x)$  be differentiable functions. Then*

$$\frac{d}{dx} \left[ \frac{f(x)}{g(x)} \right] = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$$

wherever  $g(x) \neq 0$ .

**Example 2.4.4 Derivative of tangent.** Let  $f(x) = \tan x$ . Find  $f'(x)$ .

**Solution.** Since  $\tan x = \frac{\sin x}{\cos x}$ , then we can apply the quotient rule to get the derivative of  $\tan x$ :

$$\frac{d}{dx} \tan x = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \sec^2 x.$$

The derivatives for  $\sec x, \csc x$  and  $\cot x$  may also be computed using the quotient rule and the facts that  $\frac{d}{dx} \sin x = \cos x$  and  $\frac{d}{dx} \cos x = -\sin x$ .

## 2.5 The Chain Rule

At this point, we can take derivatives of sums, differences, products and quotients of functions. However, these rules aren't very useful for differentiating functions like  $f(x) = (3x^2 + \sin x)^{50}$ . We could technically evaluate  $f'(x)$  using these rules but it would be an awful way to spend your weekend. But if we make the substitution  $u = 3x^2 + \sin x$ , then we can rewrite  $f(x)$  as  $f(u) = u^{50}$ , which is *much* easier to differentiate:  $\frac{df}{du} = 50u^{49}$ . At this step we might be tempted to say that  $f'(x) = 50u^{49} = 50(3x^2 + \sin x)^{49}$ , but this isn't quite right. To get the actual derivative we need to consider how the new variable  $u$  depends on  $x$  as well. The **Chain Rule** is what we need.

**Theorem 2.5.1 The Chain Rule.** *Let  $y = f(u)$  and  $u = g(x)$  be differentiable functions. Then*

$$\frac{dy}{dx} = \frac{dy}{du} \frac{du}{dx}.$$

*Equivalently, if we set  $F(x) = f(g(x))$  we have*

$$F'(x) = f'(g(x))g'(x).$$

**Example 2.5.2 Using the chain rule.** Let  $f(x) = (3x^2 + \sin x)^{50}$ . Find  $f'(x)$ .

**Solution.** We'll try the same trick we used before and we'll set  $u = 3x^2 + \sin x$ . Then the chain rule says that

$$\begin{aligned} f'(x) &= \frac{df}{du} \frac{du}{dx} \\ &= (50u^{49})(6x + \cos x) \\ &= 50(3x^2 + \sin x)^{49}(6x + \cos x). \end{aligned}$$

As the last example highlighted, we can use the chain rule in combination with any of the other derivative rules we know if the function we're differentiating is complicated.

**Example 2.5.3 Combining rules.** Let  $f(s) = \sqrt{\frac{s^2-1}{s^2+1}}$ . Find  $f'(s)$ .

**Solution.** We'll let  $u = \frac{s^2-1}{s^2+1}$  stand in for the "inside function." Then we have  $f(u) = \sqrt{u}$  and so

$$\begin{aligned} \frac{df}{ds} &= \frac{df}{du} \frac{du}{ds} \\ &= \frac{1}{2} u^{-\frac{1}{2}} \frac{2s(s^2+1) - 2s(s^2-1)}{(s^2+1)^2} \\ &= \frac{1}{2} \left( \frac{s^2-1}{s^2+1} \right)^{-\frac{1}{2}} \frac{2s(s^2+1) - 2s(s^2-1)}{(s^2+1)^2} \end{aligned}$$

**Example 2.5.4 Chain rule within a chain rule.** Find the slope of  $g(x) = \sin^4(\cos^3 x)$  at  $x = \pi$ .

**Solution.** We need to compute  $g'(\pi)$ . First, note that  $g(x) = [\sin(\cos^3 x)]^4$ , so let  $u = \sin(\cos^3 x)$ . Now we could try to use the chain rule right here but this would require finding  $\frac{du}{dx}$ , and  $u$  is itself a complicated function of  $x$ . So let  $v = [\cos x]^3$ , and finally let  $w = \cos x$ . Then we can say that

$$g'(x) = \frac{dg}{du} \frac{du}{dv} \frac{dv}{dw} \frac{dw}{dx}$$

$$\begin{aligned}
&= \left( \frac{d}{du} u^4 \right) \left( \frac{d}{dv} \sin v \right) \left( \frac{d}{dw} w^3 \right) \frac{d}{dx} \cos x \\
&= (4u^3)(\cos v)(3w^2)(-\sin x).
\end{aligned}$$

We could plug in what  $u, v, w$  are in terms of  $x$  and then plug in  $x = \pi$ , but it's easier to just find  $u, v, w$  at  $x = \pi$  and enter these values into the above. At  $x = \pi$  we have  $u = \sin(-1)$ ,  $v = -1$  and  $w = -1$ , so

$$g'(\pi) = 4(\sin(-1))^3 \cos(-1)(3) \cdot 0 = 0.$$

## 2.6 Implicit Differentiation

**Example 2.6.1 Derivative of an implicit function.** Consider the curve given by the equation  $x^3 - y^2 + \sin y = 3$ . Find the slope of this curve at the point  $(\sqrt[3]{3}, 0)$ .

**Solution.** We could try to solve for  $y$  and then differentiate that to find the slope, but we have a slight problem: it's impossible, at least in terms of "elementary functions". However, we can still use the chain rule to find  $y'$ , at least in terms of  $x$  and  $y$ . We just need to remember that  $y$  is a function of  $x$ . If we differentiate  $x^3 - y^2 + \sin y = 3$  with respect to  $x$ , we get

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

and so

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

So the slope of the curve at  $(\sqrt[3]{3}, 0)$  is just  $-3\sqrt[3]{3}^2$ .

The method we used to get  $\frac{dy}{dx}$  in [Example 2.6.1](#) is called **implicit differentiation**. It's extremely useful if we want to solve for  $\frac{dy}{dx}$  without first solving for  $y$ . Even if we can solve for  $y$  without too much trouble, it's often easier to find  $\frac{dy}{dx}$  implicitly as the next example shows.

**Example 2.6.2 Implicit differentiation to save algebra.** Let  $f(x) = \sqrt{1 - x^2}$ . Find  $f'(x)$ .

**Solution.** If we let  $y = f(x)$ , then  $x^2 + y^2 = 1$ . Then

$$2x + 2yy' = 0$$

which means that

$$y' = -\frac{x}{y} = -\frac{x}{\sqrt{1 - x^2}}.$$

We must also be aware of when to use appropriate derivative rules when doing implicit differentiation.

**Example 2.6.3 Chain and quotient rule.** Suppose  $y$  is defined implicitly by  $\tan(x - y) = \frac{y}{1 + x^2}$ . Find  $y'$ .

**Solution.** We start by taking the derivative with respect to  $x$  of each side of the equation:

$$\frac{d}{dx} \tan(x - y) = (1 - y') \sec^2(x - y)$$

$$\frac{d}{dx} \frac{y}{1+x^2} = \frac{y'(1+x^2) - 2xy}{(1+x^2)^2}.$$

Therefore

$$(1 - y') \sec^2(x - y) = \frac{y'(1+x^2) - 2xy}{(1+x^2)^2},$$

and we can solve this for  $y'$ :

$$(1+x^2)^2 \sec^2(x-y) + 2xy = y'(1+x^2) + (1+x^2)^2 y' \sec^2(x-y)$$

or just

$$y' = \frac{(1+x^2)^2 \sec^2(x-y) + 2xy}{(1+x^2) + (1+x^2)^2 \sec^2(x-y)}$$

**Example 2.6.4 A differential equation.** Let  $P(t)$  denote the population of the United States, where  $P$  is in millions and  $t$  is the number of years after 1990. Then the growth of  $P(t)$  can be modeled by the **differential equation**

$$P' = \frac{1}{23500} P(532 - P).$$

According to this model, it's not too hard to see that  $P'$  should be positive given the current population of the US, so the model predicts the population to increase. Is this rate of growth increasing or decreasing?

**Solution.** We need to find  $P''$ , which is the rate of change of  $P'$ . This means we need to differentiate both sides of the differential equation  $P' = \frac{1}{23500} P(532 - P)$  with respect to  $t$ :

$$P'' = \frac{1}{23500} P'(532 - P) - \frac{1}{23500} P(P').$$

The current population is about 323.1 million, so we can take  $P = 323.1$ , which also gives

$$P' = \frac{1}{23500} (323.1)(532 - 323.1) = 2.9,$$

and so

$$P'' = \frac{1}{23500} (2.9)(532 - 323.1) - \frac{1}{23500} (323.1)(2.9) < 0.$$

Hence it appears that the rate of population increase is itself decreasing, which implies that the population growth of the US is slowing down.

## 2.7 Related Rates

**Example 2.7.1 Changing volume on a sphere.** The radius of a sphere is increasing at a rate of  $4 \frac{\text{mm}}{\text{s}}$ . How fast is the volume increasing when the radius is 13 mm?

**Solution.** First, let's assign names to all of the changing quantities in this problem:

$V$  = volume

$r$  = radius

Note that we're considering  $V$  and  $r$  to be functions of time, i.e.  $V = V(t)$  and  $r = r(t)$ , where  $t$  is in seconds. Then we're given that  $r' = 4$  and we need to

find  $V'$  when  $r = 13$ . To do this, we had better find some relationship between  $V$  and  $r$ . We can find this by looking at the volume formula for a sphere:

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

This equation relates  $r$  and  $V$ , and if we take the derivatives of both sides with respect to time  $t$  (using implicit differentiation) then we get an equation relating  $V'$  and  $r'$ . So let's do that:

$$V' = 4\pi r^2 r'$$

Now we can plug in our given information to get

$$V' = 4\pi(13^2)(4) = 2704\pi.$$

So the volume is increasing at a rate of  $2704\pi \frac{\text{mm}^3}{\text{s}}$ .

Note that we *never* found what  $V(t)$  and  $r(t)$  were in [Example 2.7.1](#), but we didn't need to. All we needed to find was a relationship between the two changing quantities, i.e. the two derivatives, to answer the question.

**Example 2.7.2 A tank problem.** Water is leaking out of a conical tank at a rate of  $10000 \frac{\text{cm}}{\text{min}}$  while water is being pumped into the tank at some (unknown) constant rate. The tank has a height of 6 m and the diameter of the top is 4 m. If the water level is rising at a rate of  $20 \frac{\text{cm}}{\text{min}}$  when the height of the water is 2 m, what is the rate at which water is being poured into the tank?

**Solution.** We have a *lot* of information to process here, so we'll take things a step at a time. The changing quantities are

$$h(t) = \text{height of water}$$

$$r(t) = \text{radius of water}$$

$$V(t) = \text{volume of water}$$

where  $t$  is in minutes. We need to find the rate at which water is entering the tank, so let's call this mystery number  $C$ . Then all we know about  $C$  is that it's related to the rate that the volume is changing. In particular, we should have  $V' = C - 10000$ . So to find  $C$  we need to find  $V'$ , which means we need to set up a relationship between  $V$  and the other changing quantities to determine how exactly  $V$  is changing, using the fact that  $h' = 20$  when  $h = 200$  (after converting to centimeters).

Using the fact that the water is in a conical tank, we can use the formula for the volume of a cone to say that  $V = \frac{1}{3}\pi r^2 h$ . If we differentiate both sides with respect to  $t$ , then we get

$$V' = \frac{1}{3}\pi(2rr'h + r^2h').$$

Unfortunately, we don't have any information on  $r$  or  $r'$  that we can use, just information on  $h$  and  $h'$ . So we need to get  $r$  in terms of  $h$ . By similar triangles, we can say that  $\frac{r}{h} = \frac{200}{600} = \frac{1}{3}$ , and so  $r = \frac{h}{3}$  and likewise  $r' = \frac{h'}{3}$ . Hence

$$V' = \frac{1}{3}\pi(2rr'h + r^2h') = \frac{1}{3}\pi\left(2\frac{h^2h'}{9} + \frac{h^2h'}{9}\right)$$

which boils down to

$$V' = \frac{h^2h'}{9}\pi = \frac{800000\pi}{9}.$$



Now, *finally*, we can answer the original question. The rate that water is flowing into the tank is

$$C = V' + 10000 = \frac{800000\pi}{9} + 10000,$$

or just 289 252.68  $\frac{\text{cm}^3}{\text{min}}$ . In terms of meters this is  $0.29 \frac{\text{m}^3}{\text{min}}$ , which perhaps looks a bit more reasonable.

## 2.8 Linear Approximations

Suppose we graph the tangent line to  $y = \sqrt{x}$  at  $x = 4$ . The equation of this line will be

$$y - 2 = \frac{x - 4}{4} \quad \text{or} \quad y = 2 + \frac{x - 4}{4}.$$

If we look at the tangent line next to the graph, we'll see that the line is very close to the graph of  $y = \sqrt{x}$  if  $x$  is very close to 4. So we can use the equation of the tangent line to approximate  $y = \sqrt{x}$  when  $x$  is near 4. For example, we can say that

$$\sqrt{4.2} \approx 2 + \frac{4.2 - 4}{4} = 2.05.$$

We call this the **linear approximation** of  $\sqrt{x}$  at 4.

**Definition 2.8.1 Linear Approximation.** Let  $f(x)$  be a function that is differentiable at  $x = a$ . The linear approximation (or **linearization**) of  $f$  at  $a$  is the function  $L(x)$  given by

$$L(x) = f(a) + f'(a)(x - a).$$

The formula of the linear approximation to a function  $f$  at  $a$  is nothing more than the equation of the tangent line to  $f$  at  $a$ .

**Example 2.8.2 Linear approximation of sine.** Find the linear approximation of  $f(\theta) = \sin \theta$  at  $\theta = 0$  and use this to estimate  $\sin(-.02)$ .

**Solution.** The linear approximation is given by

$$L(\theta) = f(0) + f'(0)(\theta - 0) = \sin 0 + \cos 0(\theta - 0),$$

so  $L(\theta) = \theta$ . Therefore

$$\sin(-.02) \approx -.02$$

since  $-.02$  is very close to 0.

**Example 2.8.3 Estimating a tangent.** Estimate  $\tan(\pi + \frac{3}{49})$ .

**Solution.** We want to estimate this using a linear approximation, and we want to make sure the linear approximation is easy to set up. This means we want to base the linear approximation at a value of  $a$  that both tangent and its derivative are relatively easy to compute at, and is also close to  $\pi + \frac{3}{49}$ . So we'll pick  $a = \pi$  and find the linear approximation to  $\tan x$  at  $a = \pi$ : Set  $f(x) = \tan x$ . Then

$$\begin{aligned} L(x) &= f(a) + f'(a)(x - a) \\ &= \tan \pi + \sec^2(\pi)(x - \pi) \\ &= x - \pi. \end{aligned}$$

So

$$\tan(\pi + \frac{3}{49}) \approx (\pi + \frac{3}{49}) - \pi = \frac{3}{49}.$$

**Example 2.8.4 Estimating the population of the United States in 2020.** Use the differential equation from [Example 2.6.4](#) and the fact that the current population of the US is about 323.1 million to estimate the population of the US in 2020.

**Solution.** Recall that  $P(t)$  in [Example 2.6.4](#) represented the population of the US (in millions)  $t$  years after 1990. To estimate the population in 2020, we want to find the linearization of  $P$  at  $t = 27$ . This is given by

$$L(t) = P(27) + P'(27)(t - 27).$$

Since  $P(27) = 323.1$  and

$$P'(27) = \frac{1}{23500}P(532 - P) = \frac{1}{23500}(323.1)(532 - 323.1) = 2.9,$$

we have

$$L(t) = 323.1 + 2.9(t - 27).$$

Hence the population of the US in 2020 should be

$$L(30) = 323.1 + 2.9 \cdot 3 = 331.8,$$

or about 331.8 million people.

## Chapter 3

# Inverse Functions

Recall that a function  $f(x)$  is **invertible** if it's one-to-one. In this case, there's an inverse function  $f^{-1}(x)$  that undoes  $f(x)$ . In symbols,  $f^{-1}(f(x)) = x$  and  $f(f^{-1}(y)) = y$ .

### 3.1 Exponential Functions

Many changing quantities can be modeled by using **exponential functions**, which are functions of the form  $f(x) = a^x$  where  $a > 0$ . Exponential functions have several important properties which we list below.

**Theorem 3.1.1 Properties of Exponential Functions.** *Let  $a > 0$  and  $x, y$  be real numbers. Then*

- $a^0 = 1$  and  $a^1 = a$ .
- $a^{-x} = \frac{1}{a^x}$ .
- $a^{x+y} = a^x a^y$  and  $a^{x-y} = \frac{a^x}{a^y}$ .
- $a^{x/y} = \sqrt[y]{a^x}$  is  $y > 0$ .
- $(a^x)^y = a^{xy}$ .

Finally, if  $b > 0$  as well, then  $(ab)^x = a^x b^x$ .

**Example 3.1.2 Limit of an exponential function.** Let  $f(x) = 2^{-x}$ . Find  $\lim_{x \rightarrow \infty} f(x)$ .

**Solution.** First, note that

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

So as  $x \rightarrow \infty$ , the denominator  $2^x \rightarrow \infty$  as well. Hence the fraction as a whole should decrease to 0, and so

$$\lim_{x \rightarrow \infty} 2^{-x} = 0.$$

In general, we have the following:

$$\begin{aligned} a > 1 &\Rightarrow \lim_{x \rightarrow \infty} a^x = \infty \quad \text{and} \quad \lim_{x \rightarrow -\infty} a^x = 0 \\ a < 0 &\Rightarrow \lim_{x \rightarrow \infty} a^x = 0 \quad \text{and} \quad \lim_{x \rightarrow -\infty} a^x = \infty \end{aligned}$$

Every exponential function  $f(x) = a^x$  is continuous everywhere. In fact, they are differentiable everywhere, though we'll see more about that later. We can also say the following: if  $a > 1$  then  $a^x$  is an increasing function, and if  $0 < a < 1$  then  $a^x$  is a decreasing function. In neither case will  $a^x$  have any local maxima or minima.

Of all the different exponential functions, the most important one is the so-called **natural exponential function**, denoted  $e^x$  where

$$e = 2.71828\dots$$

This is the unique exponential function by the property that its derivative at 0 is 1. In fact, as we will see later  $e^x$  is its own derivative.

**Example 3.1.3 Exponential derivative.** Using the fact that  $\frac{d}{dx}e^x = e^x$ , compute  $\frac{d}{dx}e^{5x-x^4+\cos x}$ .

**Solution.** We'll have to use the chain rule. If we do so, we get

$$\frac{d}{dx}e^{5x-x^4+\cos x} = (5 - 4x^3 - \sin x)e^{5x-x^4+\cos x}.$$

Since  $e^x$  is its own derivative, this makes computing derivatives of functions of the form  $e^{f(x)}$  relatively straightforward:

$$\frac{d}{dx}e^{f(x)} = f'(x)e^{f(x)}.$$

Technically, we have worked with the natural exponential function before in this class. Recall from [Example 2.6.4](#) that if  $P(t)$  represents the population of the USA (in millions)  $t$  years after 1990, then we said that  $P$  satisfied the differential equation

$$\frac{dP}{dt} = \frac{1}{23500}P(532 - P).$$

If we say the population of the US in 1990 was 250 million (i.e.  $P(0) = 250$ ), then

$$P(t) = \frac{133000}{250 + 282e^{-.023t}}.$$

Now, in [Example 2.8.4](#) we used linear approximation to estimate the population of the US using this model in the year 2020, where we got 331.8 million people. If we use the above formula for  $P(t)$ , we get that the population of the US in 2020 should be about

$$P(30) = 339.8,$$

or 339.8 million people.

The Census Bureau predicted in 2014 that the population in 2020 would be 334.5 million. So our simple predictions before actually aren't too far off from what the Census Bureau is expecting. This suggests that the differential equation  $P' = \frac{1}{23500}P(532 - P)$  is a reasonable model for the growth of the US.

## 3.2 Logarithms

### 3.2.1 Inverse Function Review

Suppose we want to solve the equation  $y = f(x)$  for  $x$ . Then we can only do so unambiguously if  $f$  is an **invertible function**.

**Definition 3.2.1 Invertible Functions.** A function  $f(x)$  is invertible if and only if it is a one-to-one function. That is,  $f(x_1) = f(x_2)$  forces  $x_1 = x_2$ .

Since a function is one-to-one if and only if its graph passes the horizontal line test, the horizontal line test also gives us a useful way to check if a function is invertible.

**Example 3.2.2.** Let  $f(x) = e^x$ , the natural exponential function introduced in [Section 3.1](#). Then  $f(x)$  is an invertible function, since its graph passes the horizontal line test.

If a function  $f$  is invertible, then there exists an **inverse function**  $f^{-1}$  satisfying the following equations:

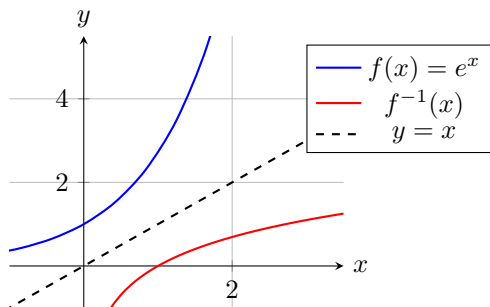
$$\begin{aligned} f(f^{-1}(y)) &= y \\ f^{-1}(f(x)) &= x. \end{aligned}$$

Essentially,  $f^{-1}$  is the rule that undoes the transformation that  $f$  applies to  $x$ . The domain of  $f^{-1}$  is the range of  $f$  and the range of  $f^{-1}$  is the domain of  $f$ . Note that  $f^{-1} \neq \frac{1}{f}$ !

The graph of an inverse function is related to the graph of the original function in a very nice way. To get the graph of  $f^{-1}$  from  $f$ , just reflect the graph of  $f$  across the line  $y = x$ .

**Example 3.2.3 Graphing the inverse of the natural exponential.** Let  $f(x) = e^x$ . Graph  $f^{-1}(x)$ .

**Solution.** If we graph  $e^x$  and then rotate the graph across the line  $y = x$ , we get



### 3.2.2 Calculus with Inverse Functions

There are two important things to notice about the graph of the inverse function in [Example 3.2.2](#): it's continuous and differentiable. This is because the same was true of the original graph, and if we think about it this should be true in general as well since graphing the inverse doesn't add any new gaps or cusps to the graph.

**Theorem 3.2.4 Continuity and Differentiability of Inverses.** Let  $f(x)$  be a continuous (respectively, differentiable) function. Suppose that  $f(x)$  is invertible, and has inverse  $f^{-1}(x)$ . Then  $f^{-1}(x)$  is continuous (respectively, differentiable).

**Example 3.2.5 Finding the derivative of an inverse function.** Let  $f(x) = \sqrt{x-1}$  and note that  $f$  is invertible. Find the derivative of  $f^{-1}$ .

**Solution.** One way to do this is just start by finding  $f^{-1}$ . So we'll set  $y = \sqrt{x-1}$ , solve for  $x$  and then switch  $x$  and  $y$ . If we do this, we get

$$y = x^2 + 1$$

so  $f^{-1}(x) = x^2 + 1$  (and note that the domain of  $f^{-1}$  is  $[0, \infty)$  since this is the range of  $f$ !). Hence

$$\frac{d}{dx}f^{-1}(x) = 2x.$$

The method used in [Example 3.2.5](#) will certainly work in simple cases. But what do we do if we can't (or don't want to) find an explicit formula for the inverse function? The following formula will help us to do this.

**Theorem 3.2.6 Derivative of an Inverse Function.** *Let  $f(x)$  be a differentiable function with inverse  $f^{-1}(x)$ . Assume that  $f^{-1}$  is itself differentiable. Then*

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

*Proof.* First, let  $y = f^{-1}(x)$ , so that we need to find  $y'$ . Then we can say that  $f(y) = x$ . Now differentiate this equation implicitly to get

$$f'(y)y' = 1$$

or just

$$y' = \frac{1}{f'(y)} = \frac{1}{f'(f^{-1}(x))}. \quad \square$$

**Example 3.2.7 Exponential inverse revisited.** Using the fact that the derivative of  $e^x$  is itself, find the derivative of its inverse.

**Solution.** Let  $f(x) = e^x$ . Then by [Theorem 3.2.6](#) we have

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

So the derivative of the inverse of  $e^x$  is  $\frac{1}{x}$ .

**Example 3.2.8 Derivative at a point.** Let  $f(x) = x^3 + 3\sin x + 2\cos x$ . Find  $(f^{-1})'(2)$

**Solution.** By [Theorem 3.2.6](#), we know that

$$(f^{-1})'(2) = \frac{1}{f'(f^{-1}(2))}.$$

By inspection,  $f(0) = 2$  which means  $f^{-1}(2) = 0$ . Since

$$f'(x) = 3x^2 + 3\cos x - 2\sin x,$$

this gives

$$(f^{-1})'(2) = \frac{1}{f'(f^{-1}(2))} = \frac{1}{f'(0)} = \frac{1}{3}.$$

### 3.2.3 Logarithms

Since exponential functions are important in mathematics and its applications, their inverses are important as well. Each exponential function  $f(x) = a^x$  ( $a > 0$ ) has an inverse function that we call the **logarithm with base  $a$** , and denote by  $f^{-1}(x) = \log_a x$ . The domain of each logarithm is  $(0, \infty)$  and the range is  $(-\infty, \infty)$ . The defining properties of the logarithm function are as follows:

$$\begin{aligned}\log_a a^x &= x \\ a^{\log_a x} &= x.\end{aligned}$$

Essentially, logarithms and exponentials cancel each other out. To put this another way,  $\log_a x$  is the *exponent* needed to turn  $a$  into  $x$ .

We can say a few things about logarithms just from looking at graphs of exponentials. Suppose that  $a > 1$ . Then  $\log_a x$  is both continuous and differentiable for all  $x > 0$ ,

$$\begin{aligned}\log_a 1 &= 0 \\ \lim_{x \rightarrow 0^+} \log_a x &= -\infty \\ \lim_{x \rightarrow \infty} \log_a x &= \infty\end{aligned}$$

and the derivative of  $\log_a x$  approaches 0 as  $x \rightarrow \infty$ .

The properties of the exponential listed in [Theorem 3.1.1](#) have corresponding properties for logarithms.

**Theorem 3.2.9 Properties of Logarithm Functions.** *Let  $a > 0$  with  $a \neq 1$ . Let  $x, y$  be positive real numbers, and let  $r$  be any real number. Then*

- $\log_a 1 = 0$  and  $\log_a a = 1$ .
- $\log_a(xy) = \log_a x + \log_a y$ .
- $\log_a \frac{x}{y} = \log_a x - \log_a y$ .
- $\log_a x^r = r \log_a x$ .

An important takeaway from [Theorem 3.2.9](#) is that logarithms turn the complicated operations of multiplication and division into the simpler operations of addition and subtraction.

Since every exponential function has a corresponding logarithm, the natural exponential function  $e^x$  has a logarithm as well. We call this inverse the **natural logarithm** and denote it by  $\ln x$ . Note that  $\frac{d}{dx} \ln x = \frac{1}{x}$  by [Example 3.2.7](#).

**Example 3.2.10 Simplifying an exponential.** Simplify  $e^{-4 \ln x}$  and  $e^{x \ln 3}$ .

**Solution.** We'll simplify by using the cancellation property  $e^{\ln y} = y$ . First, we need to put the entire exponent inside of the natural log:

$$e^{-4 \ln x} = e^{\ln x^{-4}} = x^{-4}$$

and

$$e^{x \ln 3} = 3^x.$$

[Example 3.2.10](#) shows us that *every* exponential function can be written in terms of the natural exponential:  $a^x = e^{x \ln a}$ . Similarly, every logarithm can be written in terms of the natural logarithm by using the change of base formula:

$$\log_a x = \frac{\ln x}{\ln a}.$$

### 3.3 Derivatives of Exponential and Logarithmic Functions

#### 3.3.1 Derivatives of Exponentials

We've already mentioned that  $\frac{d}{dx}e^x = e^x$ , but we've said nothing about differentiating functions such as  $3^x$  or  $(\frac{1}{\pi})^{-x}$ . However, it turns out that we can get these derivatives from the derivative of  $e^x$  without too much trouble.

**Theorem 3.3.1 Derivatives of Exponential Functions.** *Let  $a > 0$  and set  $f(x) = a^x$ . Then  $f'(x) = a^x \ln a$ .*

*Proof.* To prove this, we need to use what's available to us: the derivative of  $e^x$ . From [Example 3.2.10](#) we say that  $e^{x \ln 3} = 3^x$ . More generally, we can also say that  $e^{x \ln a} = a^x$ . Hence

$$\begin{aligned}\frac{d}{dx}a^x &= \frac{d}{dx}e^{x \ln a} \\ &= (\ln a)e^{x \ln a} \\ &= (\ln a)a^x. \quad \square\end{aligned}$$

**Example 3.3.2 Differentiating an exponential.** Let  $g(t) = 2.6^{3t - \sin t}$ . Find  $g'(t)$ .

**Solution.** If we set  $u = 3t - \sin t$ , then  $g(t) = g(u) = 2.6^u$ . By the chain rule,  $g'(t) = g'(u)\frac{du}{dt} = (\ln 2.6)2.6^u(3 - \cos t) = (\ln 2.6)2.6^{3t - \sin t}(3 - \cos t)$ .

Although we certainly want to know the derivative of  $a^x$ , or at least how to find it, most applications involving exponential functions use the natural exponential function  $e^x$  instead. The derivative of  $e^x$  is probably the most important derivative in this course.

**Example 3.3.3 Solutions of a differential equation.** Show that  $x(t) = Ae^{-t} + Bte^{-t}$  satisfies the differential equation  $x'' + 2x' + x = 0$ , where  $A, B$  are arbitrary constants.

**Solution.** We need to show that if we compute  $x'$  and  $x''$  and plug these expressions into the differential equation, this will simplify out to 0. Since

$$\begin{aligned}x' &= -Ae^{-t} + Be^{-t} - Bte^{-t} \\ x'' &= Ae^{-t} - Be^{-t} - (Be^{-t} - Bte^{-t}) \\ &= Ae^{-t} - 2Be^{-t} + Bte^{-t},\end{aligned}$$

it follows that

$$\begin{aligned}x'' + 2x' + x &= (Ae^{-t} - 2Be^{-t} + Bte^{-t}) + 2(-Ae^{-t} + Be^{-t} - Bte^{-t}) + Ae^{-t} + Bte^{-t} \\ &= 0.\end{aligned}$$

Exponential functions often appear in the solutions of differential equations, which are themselves often viewed as mathematical models of physical systems. Hence exponential functions play a significant role in predicting physical quantities, which goes a long way towards justifying their importance.



### 3.3.2 Derivatives of Logarithms

Just as we can get the derivative of every exponential function  $a^x$  just by knowing the derivative of  $e^x$ , we can get the derivative of every logarithmic function  $\log_a x$  just by knowing that  $\frac{d}{dx} \ln x = \frac{1}{x}$ .

**Theorem 3.3.4 Derivatives of Logarithmic Functions.** Let  $a > 0, a \neq 1$  and set  $f(x) = \log_a x$ . Then

$$f'(x) = \frac{1}{x \ln a}.$$

**Example 3.3.5 Differentiating nested logarithms.** Let  $f(x) = \log_2(\log_5 x)$ . Find  $f'(x)$ .

**Solution.** By the chain rule, we have

$$f'(x) = \frac{1}{(\log_5 x) \ln 2} \frac{d}{dx} \log_5 x = \frac{1}{(\log_5 x) \ln 2} \frac{1}{x \ln 5},$$

Logarithms can be used to greatly simplify derivatives involving products, division or exponentiation using a technique known as **logarithmic differentiation**.

**Example 3.3.6 A simple fraction.** Let  $f(x) = \frac{e^{-x} \cos x}{x^2 + 4x + 4}$ . Find  $f'(x)$ .

**Solution.** We can find  $f'(x)$  without resorting to logarithms, but this would require using the product, quotient and chain rules. The algebra would be awful. So we'll use logarithms instead! Set  $y = f(x)$ . Then

$$\ln y = \ln e^{-x} + \ln \cos x - \ln(x^2 + 4x + 4) = -x + \ln \cos x - 2 \ln(x + 2).$$

Now we differentiate both sides implicitly to obtain

$$\frac{y'}{y} = -1 - \frac{\sin x}{\cos x} - \frac{2}{x + 2}.$$

Hence

$$y' = y \left( -1 - \tan x - \frac{2}{x + 2} \right) = \frac{e^{-x} \cos x}{x^2 + 4x + 4} \left( -1 - \tan x - \frac{2}{x + 2} \right).$$

Logarithmic differentiation is useful for finding derivatives of expressions containing complicated quotients, products or powers.

**Example 3.3.7 A simple exponent.** Let  $h(z) = z^z$ . Find  $h'(z)$ .

**Solution.** We'll use logarithmic differentiation again to simplify  $h$  and remove the exponent. Set  $y = h(z)$ , which gives

$$\ln y = z \ln z.$$

So

$$\frac{y'}{y} = \ln z + 1,$$

which means that

$$y' = y(\ln z + 1) = z^z(\ln z + 1).$$

## 3.4 Exponential Models

### 3.4.1 Exponential Growth

Suppose we want to predict the spread of some infectious disease through a city. A reasonable, though simplistic, assumption is that the disease will spread quicker if more people are infected. In other words, we'll assume that the rate at which people are infected is proportional to the number of people infected. If we let  $N(t)$  denote the number of people infected at time  $t$ , then we're basically saying that

$$\frac{dN}{dt} = kN(t)$$

for some constant  $k > 0$ .

This differential equation is a **mathematical model** that allows us to predict the growth of  $N(t)$ . We can actually use it to get an expression for  $N(t)$ . This differential equation is saying that  $N(t)$  is a function whose derivative looks quite a bit like  $N(t)$ . In fact, the solution of this differential equation is  $N(t) = Ce^{kt}$ , where  $C = N(0)$  is a constant that represents the number of people infected at time  $t = 0$ . This is an example of **exponential growth**, and the proportionality constant  $k$  is called the **relative growth rate**.

**Example 3.4.1 Modeling an Outbreak.** A game of Humans vs. Zombies breaks out at WVWC. When the game starts, there is just one zombie. After one hour, there are 3 zombies. How many zombies will there be after three hours assuming that the relative growth rate is constant?

**Solution.** Let  $N(t)$  denote the number of zombies  $t$  hours after the game starts. Then we're given that  $N(0) = 1$  and  $N(1) = 3$ . We need to find  $N(3)$ .

Since the relative growth rate is assumed to be constant, we can say that  $N' = kN$  for some constant  $k$ . As we saw before, the solution of this differential equation is given by  $N(t) = N(0)e^{kt} = e^{kt}$ . So if we want to find  $N(3)$ , then we need to figure out what  $k$  is.

We can do this by using the fact that  $N(1) = 3$ . If we plug this into  $N(t) = e^{kt}$ , we get

$$3 = e^k \Rightarrow k = \ln 3.$$

So

$$N(t) = e^{(\ln 3)t} = 3^t.$$

Hence there will be  $N(3) = 3^3 = 27$  zombies after three hours.

### 3.4.2 Exponential Decay

Closely related to the concept of exponential growth is that of **exponential decay**. A quantity  $M(t)$  undergoes exponential decay if it satisfies the differential equation

$$\frac{dM}{dt} = kM,$$

where  $k < 0$ . Every solution of this differential equation looks like  $M(t) = Ce^{kt} = M(0)e^{kt}$ .

We often write the decay constant  $k$  as positive, and then rewrite the ODE for  $M(t)$  as  $\frac{dM}{dt} = -kM$ . The solution becomes  $M(t) = M(0)e^{-kt}$ . This has the effect of highlighting the negative growth rate inherent to this system.

Perhaps the most important example of exponential decay is that of radioactive decay. If  $m(t)$  represents the mass of some radioactive substance, then experiments show that substance decays at a rate proportional to the amount of the substance remaining. In symbols,  $\frac{dm}{dt} = km$  where  $k < 0$ . If we let  $m_0 = m(0)$  denote the initial mass of the substance, then we can say that  $m(t) = m_0 e^{kt}$ . The rate of decay is often expressed in terms of **half-life**: the amount of time it will take for precisely half of the mass to decay.

**Example 3.4.2 Decay from Half-Life.** A radioactive substance has a half-life of 28 days. If we start with a 50 mg sample, how much of the mass will remain after  $t$  days?

**Solution.** If we let  $m(t)$  denote the mass of the sample at time  $t$ , and let  $t = 0$  denote the first day we have the sample, then

$$m(t) = m_0 e^{kt} = 50e^{kt}.$$

We still need to find  $k$ , but we can do this using the fact that the half-life is 28 days. This means that

$$m(28) = \frac{m(0)}{2} = 25.$$

So

$$25 = 50e^{28k} \Rightarrow \ln \frac{1}{2} = 28k \Rightarrow k = -\frac{\ln 2}{28}.$$

Hence

$$m(t) = 50e^{-\frac{\ln 2}{28}t} = 50 \cdot 2^{-\frac{t}{28}}.$$

### 3.5 Inverse Trigonometric Functions

Recall that  $\sin x$  is not a one-to-one function since it fails the horizontal line test. Hence it has no inverse function. However, we can *restrict* the domain of  $\sin x$  so that what's left over passes the horizontal line test. For example, if we define  $f(x) = \sin x$  but restrict the domain to the interval  $[-\frac{\pi}{2}, \frac{\pi}{2}]$ , then  $f(x)$  is one-to-one and so has an inverse function. We call this function **inverse sine** or **arcsine**.

**Definition 3.5.1 Inverse Sine.** Let  $f(x) = \sin x$  and have domain  $[-\frac{\pi}{2}, \frac{\pi}{2}]$ . Then  $f^{-1}(x)$  is called the inverse sine or arcsine of  $x$ , and is denoted by  $\sin^{-1}x$  or  $\arcsin x$ . The domain is the interval  $[-1, 1]$  and the range is  $[-\frac{\pi}{2}, \frac{\pi}{2}]$ .

We can think of  $\sin^{-1}x = \arcsin x$  as the angle required to turn sine into  $x$ .

Note that  $\sin^{-1}x$  does *not* mean  $\frac{1}{\sin x}$ , which is  $\csc x$ . It's an unfortunate, though standard, choice of notation.

**Example 3.5.2 Finding inverse sine.** Find  $\sin^{-1}(\frac{\sqrt{3}}{2})$ .

**Solution.** This is the angle between  $-\frac{\pi}{2}$  and  $\frac{\pi}{2}$  that turns sine into  $\frac{\sqrt{3}}{2}$ . So  $\sin^{-1}(\frac{\sqrt{3}}{2}) = \frac{\pi}{3}$ .

**Example 3.5.3 Simplifying cosine and inverse sine.** Simplify  $\cos(\arcsin x)$ , where  $-1 \leq x \leq 1$ .

**Solution.** By definition,  $\sin(\arcsin x) = x$  as long as  $-1 \leq x \leq 1$ . So we want to try to rewrite  $\cos(\arcsin x)$  to make use of this cancellation property. We can do this using the Pythagorean identity  $\cos \theta = \pm \sqrt{1 - \sin^2 \theta}$ . Hence

$$\cos(\arcsin x) = \sqrt{1 - (\sin(\arcsin x))^2} = \sqrt{1 - x^2}$$

whenever  $-1 \leq x \leq 1$ .

Since  $\sin x$  is differentiable, this means  $\arcsin x$  is also differentiable.

**Theorem 3.5.4 Derivative of Inverse Sine.** Let  $-1 < x < 1$ . Then  $\frac{d}{dx} \sin^{-1} x = \frac{1}{\sqrt{1-x^2}}$ .

*Proof.* If we set  $y = \sin^{-1} x$ , then we get  $\sin y = x$ . Differentiating this gives  $y' \cos y = 1$  or just

$$y' = \frac{1}{\cos y} = \frac{1}{\sqrt{1-x^2}}. \quad \square$$

We can also define the **inverse cosine function** in much the same way as the inverse sine function. The idea once again is to restrict the domain of  $\cos x$  to  $[0, \pi]$  to get an invertible function. We call the inverse of this function the inverse cosine, or arccosine, function. We denote this function by  $\cos^{-1} x$  or  $\arccos x$ .

**Example 3.5.5 Inverse cosine value.** Find  $\csc(\cos^{-1}(-\frac{1}{2}))$ .

**Solution.** First, note that  $\cos^{-1}(-\frac{1}{2}) = \frac{4\pi}{3}$ . So

$$\csc(\cos^{-1}(-\frac{1}{2})) = \csc(\frac{4\pi}{3}) = \frac{2}{\sqrt{3}}.$$

Inverse cosine behaves in much the same way as inverse sine. However, this function won't be as useful to us as the inverse sine function or the next function we will look at: the inverse tangent.

If we restrict the domain of  $\tan x$  to  $(-\frac{\pi}{2}, \frac{\pi}{2})$ , then we get an invertible function. We call the inverse of this function the **inverse tangent** or **arctangent**, and denote it by  $\tan^{-1} x$  or  $\arctan x$ . The domain of  $\arctan x$  is  $(-\infty, \infty)$  and the range is  $(-\frac{\pi}{2}, \frac{\pi}{2})$ .  $\tan^{-1} x$  can be thought of as the angle between  $-\frac{\pi}{2}$  and  $\frac{\pi}{2}$  that makes tangent equal to  $x$ . It satisfies the following cancellation properties:  $\tan(\tan^{-1} x) = x$  for all  $x$  and  $\tan^{-1}(\tan y) = y$  for all  $y$  in  $(-\frac{\pi}{2}, \frac{\pi}{2})$ .

**Example 3.5.6 Simplifying an inverse tangent.** Simplify  $\sin(2 \arctan x)$  using the formula  $\sin 2x = 2 \sin x \cos x$ .

**Solution.** If we use the double-angle formula given, we get

$$\sin(2 \arctan x) = 2 \sin(\arctan x) \cos(\arctan x).$$

So we just need to find  $\sin(\arctan x)$  and  $\cos(\arctan x)$ .

Set  $\theta = \arctan x$ , so that  $\tan \theta = x$ . Then we can find  $\sin \theta$  and  $\cos \theta$  using triangles, which gives

$$\begin{aligned} \sin \theta &= \frac{x}{\sqrt{1+x^2}} \\ \cos \theta &= \frac{1}{\sqrt{1+x^2}} \end{aligned}$$

Hence

$$\begin{aligned}\sin(2 \arctan x) &= \sin 2\theta \\ &= 2 \sin \theta \cos \theta \\ &= \frac{2x}{1+x^2}.\end{aligned}$$

**Example 3.5.7 Limit of inverse tangent.** Determine  $\lim_{x \rightarrow \infty} \tan^{-1} x$ .

**Solution.** Recall that  $\tan^{-1} x$  is the angle between  $-\frac{\pi}{2}$  and  $\frac{\pi}{2}$  for which tangent is equal to  $x$ . So finding  $\lim_{x \rightarrow \infty} \tan^{-1} x$  is equivalent to finding what angle between  $-\frac{\pi}{2}$  and  $\frac{\pi}{2}$  we have to approach in order for tangent to blow up to  $\infty$ . Either from looking at a graph or by using the definition of tangent, we see that the angle we need to approach is exactly  $\frac{\pi}{2}$ . Hence  $\lim_{x \rightarrow \infty} \tan^{-1} x = \frac{\pi}{2}$ .

**Theorem 3.5.8 Derivative of Inverse Tangent.** *The derivative of  $\tan^{-1} x$  is  $\frac{1}{1+x^2}$ .*

*Proof.* We can prove this the same way we proved that  $\frac{d}{dx} \arcsin x = \frac{1}{\sqrt{1-x^2}}$ . All we need to do is to set  $y = \tan^{-1} x$ , and then find  $y'$  by implicit differentiation.  $\square$

**Example 3.5.9 Tangent half-angle substitution.** An important substitution in integral calculus is the **tangent half-angle substitution** defined by  $t = \tan \frac{x}{2}$ . Use this equation to find  $\frac{dx}{dt}$ .

**Solution.** We can find  $x'$  implicitly, but we can also solve for  $x$  to get

$$x = 2 \tan^{-1} t.$$

Therefore

$$\frac{dx}{dt} = 2 \frac{d}{dt} \tan^{-1} t = \frac{2}{1+t^2}.$$

## 3.6 L'Hospital's Rule

Consider the important limit  $\lim_{x \rightarrow 0} \frac{\sin x}{x}$ . If we try to plug in  $x = 0$ , we get  $\frac{0}{0}$ , which is undefined. However, we can prove using geometry that the limit is 1. As another example, consider  $\lim_{x \rightarrow \infty} \frac{1-x^3}{4x^3}$ . Once again, if we try to plug the limit in we get an expression of the form  $\frac{-\infty}{\infty}$ . However, the limit is just  $-\frac{1}{4}$ , which we can find using algebra.

Limits of the form  $\frac{0}{0}$  or  $\frac{\infty}{\infty}$  are known as **indeterminate forms**. There is no restriction whatsoever on what value a limit involving an indeterminate form may take, or even that it has to exist at all.

**Example 3.6.1 Different indeterminate forms.** Find  $\frac{\infty}{\infty}$  indeterminate forms that, respectively, evaluate to 1, 0 and do not exist.

**Solution.** Let  $f(x) = \frac{1-x}{3-x}$ ,  $g(x) = \frac{1+x^4}{1+x^3}$  and  $h(x) = \frac{\ln x}{\cot x}$ . Then  $\lim_{x \rightarrow \infty} f(x)$ ,  $\lim_{x \rightarrow \infty} g(x)$  and  $\lim_{x \rightarrow 0^+} h(x)$  are all  $\frac{\infty}{\infty}$  indeterminate forms. The first evaluates to 1, the second does not exist (it's  $\infty$ ) while the third appears to be equal to 0.

The goal of this section is to determine a method that can help us evaluate limits involving indeterminate forms. This method is called **L'Hospital's Rule**.

**Theorem 3.6.2 L'Hospital's Rule.** Let  $f(x)$  and  $g(x)$  be differentiable functions. If  $\lim_{x \rightarrow a} \frac{f(x)}{g(x)}$  is either one of the indeterminate forms  $\frac{0}{0}$  or  $\frac{\infty}{\infty}$ , then

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}$$

if the limit on the right exists or is  $\pm\infty$ .

It's important to note that L'Hospital's Rule does not necessarily ask you to use the quotient rule!

**Example 3.6.3 Using L'Hospital's Rule.** Let  $h(x) = \frac{\ln x}{\cot x}$ . Find  $\lim_{x \rightarrow 0^+} h(x)$ .

**Solution.** We saw in [Example 3.6.1](#) that this limit gives us the indeterminate form  $\frac{\infty}{\infty}$ , so L'Hospital's Rule applies. Hence

$$\begin{aligned} \lim_{x \rightarrow 0^+} \frac{\ln x}{\cot x} &= \lim_{x \rightarrow 0^+} \frac{1/x}{-\csc^2 x} \\ &= \lim_{x \rightarrow 0^+} \left( -\frac{\sin x}{x} \sin x \right) \\ &= 1 \cdot 0 \end{aligned}$$

and so the limit is indeed 0.

L'Hospital's Rule also applies for limits approaching  $\pm\infty$ .

**Example 3.6.4 Exponential and polynomial growth.** Let  $f(x) = ax^2 + bx + c$  and let  $g(x) = e^x$ , where  $a, b, c$  are arbitrary constants. Find  $\lim_{x \rightarrow \infty} \frac{ax^2 + bx + c}{e^x}$ .

**Solution.** This is another example of a  $\frac{\infty}{\infty}$  indeterminate form, and so L'Hospital's Rule applies. If we use it, we get

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{ax^2 + bx + c}{e^x} &= \lim_{x \rightarrow \infty} \frac{2ax + b}{e^x} \\ &= \lim_{x \rightarrow \infty} \frac{2a}{e^x} \\ &= 0. \end{aligned}$$

In other words, the exponential function grows faster than *any* quadratic function.

**Example 3.6.5 Another limit involving exponentials.** Find  $\lim_{t \rightarrow 1^+} \frac{9^{t-1} - 4^{t-1}}{t-1}$ .

**Solution.** This is a  $\frac{0}{0}$  indeterminate form, so L'Hospital's Rule applies. We get

$$\begin{aligned} \lim_{t \rightarrow 1^+} \frac{9^{t-1} - 4^{t-1}}{t-1} &= \lim_{t \rightarrow 1^+} \frac{9^{t-1} \ln 9 - 4^{t-1} \ln 4}{1} \\ &= \ln \frac{9}{4}. \end{aligned}$$

L'Hospital's Rule only applies *directly* to the indeterminate forms  $\frac{0}{0}$  and  $\frac{\infty}{\infty}$ , but these are not the only problems when L'Hospital's Rule proves useful.

**Example 3.6.6 A different indeterminate form.** Find  $\lim_{x \rightarrow \infty} xe^{-3x-x^2}$ .

**Solution.** If we try to evaluate the limit, we get the expression  $\infty \cdot 0$ . This is another indeterminate form, but is not one that we can apply L'Hospital's Rule to without doing some algebra first. We can write

$$\lim_{x \rightarrow \infty} xe^{-3x-x^2} = \lim_{x \rightarrow \infty} \frac{x}{e^{3x+x^2}}$$

which is a  $\frac{\infty}{\infty}$  indeterminate form. We *can* use L'Hospital's on this expression!

If we do so, we get

$$\begin{aligned}\lim_{x \rightarrow \infty} \frac{x}{e^{3x+x^2}} &= \lim_{x \rightarrow \infty} \frac{1}{(3+2x)e^{3x+x^2}} \\ &= 0.\end{aligned}$$

So  $\lim_{x \rightarrow \infty} x e^{-3x-x^2} = 0$ .

**Example 3.6.6** shows that  $0 \cdot \infty$  indeterminate forms can be dealt with by rewriting them in the form  $\frac{0}{0}$  or  $\frac{\infty}{\infty}$  and then applying L'Hospital's Rule. It's trickier, but we can also deal with  $\infty - \infty$  indeterminate forms by rewriting them in this way.

**Example 3.6.7  $\infty - \infty$  indeterminate form.** Find  $\lim_{x \rightarrow \frac{\pi}{2}^-} (\sec x - \tan x)$ .

**Solution.** As  $x \rightarrow \frac{\pi}{2}^-$ ,  $\sec x \rightarrow \infty$  and  $\tan x \rightarrow \infty$ . So this limit is the indeterminate form  $\infty - \infty$ . We can't use L'Hospital's Rule yet, but we can try to rewrite this limit as a  $\frac{0}{0}$  or  $\frac{\infty}{\infty}$  indeterminate form. We'll try to do this by replacing  $\sec x, \tan x$  with  $\sin x, \cos x$  to get

$$\begin{aligned}\lim_{x \rightarrow \frac{\pi}{2}^-} (\sec x - \tan x) &= \lim_{x \rightarrow \frac{\pi}{2}^-} \frac{1 - \sin x}{\cos x} \\ &= \lim_{x \rightarrow \frac{\pi}{2}^-} \frac{-\cos x}{-\sin x} \\ &= 0.\end{aligned}$$

**Example 3.6.8 Limit involving radicals.** Find  $\lim_{x \rightarrow \infty} (\sqrt{2x^2 - x} - \sqrt{2x^2 + 1})$ .

**Solution.** This is another  $\infty - \infty$  form, so we'll try to rewrite it into a form we can use L'Hospital's Rule on. The trick here is to factor an  $x$  out so we can get a  $0 \cdot \infty$  form, which we've already seen how to handle:

$$\begin{aligned}\lim_{x \rightarrow \infty} (\sqrt{2x^2 - x} - \sqrt{2x^2 + 1}) &= \lim_{x \rightarrow \infty} (x\sqrt{2 - x^{-1}} - x\sqrt{2 + x^{-2}}) \\ &= \lim_{x \rightarrow \infty} \frac{\sqrt{2 - x^{-1}} - \sqrt{2 + x^{-2}}}{\frac{1}{x}} \\ &= \lim_{x \rightarrow \infty} \frac{\frac{1}{2}(2 - x^{-1})^{-1/2}(x^{-2}) - \frac{1}{2}(2 + x^{-2})^{-1/2}(-2x^{-3})}{-x^{-2}} \\ &= \lim_{x \rightarrow \infty} \frac{\frac{1}{2}(2 - x^{-1})^{-1/2} - \frac{1}{2}(2 + x^{-2})^{-1/2}(-2x^{-1})}{-1} \\ &= -\frac{1}{2\sqrt{2}}.\end{aligned}$$

There are three other indeterminate forms that L'Hospital's Rule can help us with (after some algebra):  $1^\infty$ ,  $\infty^0$  and  $0^0$ . All of these can be found by first using logarithms.

**Example 3.6.9 A natural limit.** Find  $\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x$ .

**Solution.** This limit is a  $1^\infty$  indeterminate form. We'll try taking logarithms to rewrite it as a  $0 \cdot \infty$  form, so set  $y = \lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x$ . Then

$$\ln y = \lim_{x \rightarrow \infty} x \ln \left(1 + \frac{1}{x}\right)$$

is a  $0 \cdot \infty$  form. We can use L'Hospital's by rewriting this limit as a  $\frac{0}{0}$  form or  $\frac{\infty}{\infty}$ . Either way, we get  $\ln y = 1$  and so the original limit is  $y = e$ .





## Chapter 4

# Applications of Differentiation

If a function is differentiable, then its derivative gives us information about how that function is changing. This information is important in many different applications of calculus.

### 4.1 Extreme Values of Functions

It's often of interest to determine how large or small some quantity can get.

**Definition 4.1.1 Absolute Extrema.** Let  $f(x)$  be a function defined on some domain  $D$ . Let  $c$  be in  $D$ . Then

- $f(c)$  is an **absolute maximum** of  $f$  on  $D$  if  $f(c) \geq f(x)$  for all  $x$  in  $D$ .
- $f(c)$  is an **absolute minimum** of  $f$  on  $D$  if  $f(c) \leq f(x)$  for all  $x$  in  $D$ .

These values, if they exist, are **extreme values**.

**Example 4.1.2.** List the extreme values, if any, of the following functions:

1.  $f(x) = e^{-x}$ .
2.  $g(x) = x^2$ .
3.  $h(x) = \sin x$ .

Some functions may not have extreme values, but they could have values that are smaller or larger than all other values of the function "nearby".

**Definition 4.1.3 Local extrema.** Let  $f(x)$  be a function and let  $c$  be in the domain of  $f$ . Then

- $f(c)$  is a **local maximum** if  $f(c) \geq f(x)$  for all  $x$  near  $c$ .
- $f(c)$  is a **local minimum** if  $f(c) \leq f(x)$  for all  $x$  near  $c$ .

Roughly, local maxima correspond to "hilltops" whereas local minima correspond to "valleys" in the graph of  $f(x)$ .

In general, local extrema and absolute extrema can be different. However, the following theorem does provide a relationship between the two on *closed, bounded* intervals.

**Theorem 4.1.4 Extreme Value Theorem.** *Let  $f(x)$  be a continuous function defined on the interval  $[a, b]$ . Then  $f(x)$  has both an absolute maximum and absolute minimum on  $[a, b]$ . Furthermore, these values occur at either an endpoint or a local extrema.*

What Theorem 10.7.7 tells us is that we can find the extreme values of any continuous function defined on a closed, bounded interval by just checking the function at the endpoint and at its local extrema.

**Example 4.1.5 Finding extreme values.** Let  $f(x) = (x^2 - 1)^2$ . Given that  $f$  has local extrema at  $x = -1, 0, 1$ , find the extreme values of  $f$  on the interval  $[\frac{1}{2}, 6]$ .

**Solution.** Theorem 10.7.7 tells us that we can find the extreme values by checking local extrema and the endpoints of our interval. Since  $x = 1$  is the only local extreme value inside of our interval, that's the only one we need to check. We have

$$f(\frac{1}{2}) = \frac{9}{16}, f(1) = 0 \quad \text{and} \quad f(6) = 35^2.$$

So the absolute minimum of  $f$  on  $[\frac{1}{2}, 6]$  is 0 and the absolute maximum is  $35^2$ .

So if we can find where a function has local extrema, then finding absolute extrema won't be too much more difficult. Thankfully, this is relatively straightforward if the function is differentiable.

**Theorem 4.1.6 Fermat's Theorem.** *Let  $f(x)$  be a function and let  $c$  be a local extreme value of  $f$ . If  $f'(c)$  exists, then  $f'(c) = 0$ .*

So finding local extrema of  $f(x)$  often amounts to finding where  $f'(x) = 0$ , i.e., where it has a root. However, we also need to worry about where  $f'(x)$  doesn't exist (just think about  $y = |x|$ ). This leads to the following definition.

**Definition 4.1.7 Critical Points.** Let  $f(x)$  be a differentiable function. The **critical points** of  $f(x)$  are the points where  $f'(x) = 0$  or  $f'(x)$  does not exist.

Our method for finding extreme values on a closed interval will proceed as follows: find all the critical points, and then compare the values of our function at the critical points and the endpoints of the interval.

**Example 4.1.8 Extreme values of tangent.** Let  $f(\theta) = \frac{4}{3}\theta - \tan \theta$ . Find the extreme values of  $f$  on the interval  $[-\frac{\pi}{3}, \frac{\pi}{3}]$ .

**Solution.** We need to find the critical points first. We have

$$f'(\theta) = \frac{4}{3} - \sec^2 \theta.$$

This is 0 if  $\cos \theta = \frac{\sqrt{3}}{2}$ , which occurs in our interval only if  $\theta = \pm \frac{\pi}{6}$ . So we need to check  $f$  at  $\pm \frac{\pi}{3}, \pm \frac{\pi}{6}$ .

**Example 4.1.9 Dealing with a cusp.** Let  $g(x) = x^{1/3} - x^{4/3}$ . Find the absolute extrema of  $g$  on  $[-1, 1]$ .

**Solution.** First, find the critical points of  $g$ :

$$g'(x) = \frac{1}{3}x^{-\frac{2}{3}} - \frac{4}{3}x^{\frac{1}{3}} = \frac{1}{3}x^{-\frac{2}{3}}(1 - 4x).$$

This is 0 at  $x = \frac{1}{4}$  and undefined at 0, so our critical points are  $x = 0, \frac{1}{4}$ . So to find the absolute extrema we need to check  $g$  at  $x = -1, 0, \frac{1}{4}, 1$ . Doing so, we see that the absolute minimum value is  $-2$  at  $x = -1$  and the absolute maximum is  $(\frac{1}{4})^{1/3} - (\frac{1}{4})^{4/3}$  at  $x = \frac{1}{4}$ .

## 4.2 The Mean Value Theorem

The linear approximations we came up with in [Section 2.8](#) are useful for estimating complicated functions with simpler, linear models. In essence, we use the derivative of a function to tell us how much to change a given value of the function in order to estimate that function. There is one problem with this approach, at least currently. We have good reason to suspect that our approximations are close to the exact values in certain circumstances, but we don't know how close. The goal of this section is to derive an estimate for the derivative that can help us to find more precise approximations.

### 4.2.1 Rolle's Theorem

We start with a theorem that is very reasonable, at least geometrically. Let  $f(x)$  be a differentiable function on some interval  $[a, b]$ , and suppose  $f(a) = f(b)$ . Then this means the graph of  $f$  must "turn around" at some point in  $[a, b]$ , i.e., there is a local maximum or minimum contained within the interval. Combining this observation with [Theorem 4.1.6](#) gives us **Rolle's Theorem**.

**Theorem 4.2.1 Rolle's Theorem.** *Let  $f(x)$  be a differentiable function on  $[a, b]$  and suppose  $f(a) = f(b)$ . Then there exists some number  $c$  in  $(a, b)$  such that  $f'(c) = 0$ .*

[Theorem 4.2.1](#) is an example of an **existence theorem**. It tells us nothing about how to find the number  $c$ , or how many such  $c$  can exist. It only tells us that there is at least one number  $c$  in  $(a, b)$  for which  $f'(c) = 0$ . This may not seem very useful, but existence theorems can be quite powerful in mathematics.

**Example 4.2.2 Rolle's Theorem and Traffic.** A car enters a highway going at  $45 \frac{\text{mi}}{\text{h}}$  and leaves the highway going at the same speed. Was the car's acceleration ever 0?

**Solution.** The speed of the car can be represented by a velocity function  $v(t)$ . If we assume that the car entered the highway at time  $t_{\text{in}}$  and left at some future time  $t_{\text{out}}$ , then we know that

$$v(t_{\text{in}}) = v(t_{\text{out}}) = 45.$$

By [Theorem 4.2.1](#), there must be some time  $t_1$  between  $t_{\text{in}}$  and  $t_{\text{out}}$  for which  $v'(t_1) = 0$ . Since  $v'$  is exactly the acceleration, we know that the car had to stop accelerating somewhere on the highway.

### 4.2.2 The Mean Value Theorem

[Rolle's Theorem](#) is powerful because it has very general conditions for its use. However, it does require the function in question to take on the same values at the endpoints of  $[a, b]$ , and this is a condition we'd like to try to relax. However, we'll try to be clever about this and use [Rolle's Theorem](#) to do most of the heavy lifting for us.

If we imagine graphing some differentiable function  $f(x)$  on some interval  $[a, b]$ , but  $f(a) \neq f(b)$ , then we can't apply [Rolle's Theorem](#). But maybe we can adjust it just a little bit so that we can? In particular, the only reason we can't use [Theorem 4.2.1](#) is that  $f(a) \neq f(b)$ . But if we subtract the line through these points from  $f(x)$ , we should get a new function for which [Theorem 4.2.1](#) applies.

The line through these points has equation

$$y = f(a) + \frac{f(b) - f(a)}{b - a}(x - a).$$

So define

$$g(x) = f(x) - y = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Then

$$g(a) = g(b) = 0,$$

So [Theorem 4.2.1](#) *does* apply to this function. Hence there exists some number  $c$  between  $a$  and  $b$  for which  $g'(c) = 0$ .

But

$$g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a},$$

and if this is 0 then we can solve  $g'(c) = 0$  for  $f'(c)$  to get

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

This gives us the **Mean Value Theorem**.

**Theorem 4.2.3 Mean Value Theorem.** *Let  $f(x)$  be differentiable on some interval  $[a, b]$ . Then there exists a number  $c$  in  $(a, b)$  such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a},$$

or equivalently

$$f(b) - f(a) = f'(c)(b - a).$$

The [Mean Value Theorem](#) essentially says that there is at least one point inside of  $(a, b)$  for which the slope  $f'(c)$  at that point matches the "average slope"  $\frac{f(b) - f(a)}{b - a}$ . Like [Rolle's Theorem](#), this is an existence theorem. However, it's slightly more general, and so is applicable in more situations. It's also useful in deriving error estimates.

**Example 4.2.4 Estimating error in a linear approximation.** Let  $f(x) = \sqrt{x}$  and let  $L(x)$  denote the linear approximation to  $f(x)$  at  $a = 9$ . Find the largest possible error between  $f(x)$  and  $L(x)$  on the interval  $[9, 16]$ .

**Solution.** First, we should find  $L(x)$ , which from [Section 2.8](#) is just

$$L(x) = 3 + \frac{1}{6}(x - 9).$$

The error we need to estimate is  $|f(x) - L(x)|$ , for  $x$  in  $[9, 16]$ . So we'll set  $g(x) = f(x) - L(x)$  and let  $x$  be some number in  $[9, 16]$ . By [Theorem 4.2.3](#), there exists some number  $c$  in  $(9, x)$  such that

$$g(x) - g(9) = g'(c)(x - 9).$$

Now,  $g(9) = 0$  and  $g'(c) = \frac{1}{2\sqrt{c}} - \frac{1}{6}$ . So the above equation becomes

$$f(x) - L(x) = \left( \frac{1}{2\sqrt{c}} - \frac{1}{6} \right) (x - 9).$$

Now, we don't know anything about  $c$  aside from the fact that it lives in  $(9, 16)$ . However, we do have an expression for  $|f(x) - L(x)|$  now:

$$|f(x) - L(x)| = \left| \left( \frac{1}{2\sqrt{c}} - \frac{1}{6} \right) (x - 9) \right|$$

$$\leq \left(\frac{1}{6} - \frac{1}{8}\right) 7.$$

So if  $x$  is in  $[9, 16]$ , then the error  $|f(x) - L(x)|$  is at most  $\frac{7}{24}$ . In other words,

$$L(x) - \frac{7}{24} \leq f(x) \leq L(x) + \frac{7}{24}.$$

Perhaps the most important example of [Theorem 4.2.3](#) in this course will be the following theorem.

**Theorem 4.2.5 Zero Derivative Theorem.** *Let  $f(x)$  be differentiable on  $(a, b)$ . If  $f'(x) = 0$  for all  $x$  in  $(a, b)$ , then  $f(x)$  has to be a constant.*

*Proof.* The idea of the proof is to use [Theorem 4.2.3](#) to show that  $f(x_2) - f(x_1) = 0$  for any points  $x_1, x_2$  in  $(a, b)$ . Now, by [Theorem 4.2.3](#) we have

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1)$$

for some  $c$  in  $(a, b)$ . However,  $f'(c) = 0$  by assumption, and so  $f(x_2) - f(x_1) = 0$ . In other words,  $f(x_2) = f(x_1)$ . Hence  $f$  must be constant.  $\square$

## 4.3 Derivatives and Graphs

Recall that if  $f(x)$  is a function and  $f'(x)$  exists, then  $f'(c) > 0$  for some  $c$  means that  $f$  is increasing at  $c$ , while  $f'(c) < 0$  means that  $f$  is decreasing at  $c$ . Now, suppose that  $c$  is a critical point of  $f(x)$ . Then  $f'(c) = 0$ . If  $f'$  changes sign from negative to positive, then  $f'(c)$  is a local minimum. Conversely, if  $f'$  changes sign from positive to negative, then  $f'(c)$  is a local maximum. This is the **first derivative test**.

**Example 4.3.1 Local maxima and minima using the first derivative test.** Let  $f(x) = x^4 e^{-x}$ . Find where  $f$  is increasing, decreasing, and any local maxima or minima.

**Solution.** We can answer all of these questions by setting up a sign chart for

$$f'(x) = x^3 e^{-x} (4 - x).$$

The critical points are  $x = 0, 4$ , and  $f' < 0$  on  $(-\infty, 0) \cup (4, \infty)$  and  $f' > 0$  on  $(0, 4)$ . So  $f$  is decreasing on the first set of intervals, increasing on  $(0, 4)$ ,  $f$  has a local maximum at  $x = 4$  and a local minimum at  $x = 0$ .

**Example 4.3.2 First derivative test and a discontinuous function.**

Find any local maxima or minima of  $G(x) = \frac{x^2}{x-1}$ .

**Solution.** We need to find the critical points, which means we need to find  $G'(x)$ :

$$G'(x) = \frac{(x-1)(2x) - x^2}{(x-1)^2} = \frac{x(x-2)}{(x-1)^2}.$$

So the critical points are  $x = 0, 1, 2$ . Note that  $x = 1$  *cannot* be a local extreme value of  $G$  since it's not in the domain of  $G$ . However, we still need to include it in our sign chart. If we do so, we find that  $G$  has a local minimum at  $x = 2$  and a local maximum at  $x = 0$  by the first derivative test.

One benefit of the first derivative test is that we only need to compute first derivatives to use it. However, if a function has a second derivative, then it's often easier to use the concavity of the graph at a critical point to determine

whether a critical point is a local maximum or minimum. In particular, if  $f(x)$  is a function and  $f''(x)$  is continuous near the critical point  $c$  (so  $f'(c) = 0$ ), then  $f(c)$  is a local minimum if  $f''(c) > 0$  and  $f(c)$  is a local maximum if  $f''(c) < 0$ . This is the **second derivative test**.

**Example 4.3.3 Using the second derivative test.** Find any local extrema of  $h(x) = 5x^3 - 3x^5$ .

**Solution.** First, we find the critical points:

$$h'(x) = 15x^2 - 15x^4 = 0$$

forces  $x = 0, \pm 1$ . Now we check these critical points in  $h''(x) = 30x - 60x^3$ :

$$\begin{aligned} h''(0) &= 0 \\ h''(-1) &= 30 \\ h''(1) &= -30. \end{aligned}$$

So  $h$  has a local minimum at  $-1$  and a local maximum at  $1$ .

The point  $x = 0$  in [Example 4.3.3](#) is an example of an **inflection point** of  $h(x)$ : a place where the concavity of  $h$  changes, or equivalently a point where the second derivative changes sign.

## 4.4 Optimization Problems

Now that we can use derivatives to find local maxima and minima, we can solve **optimization problems**.

**Example 4.4.1 Minimizing a product.** Find two numbers that differ by 100 and whose product is a minimum.

**Solution.** Call our numbers  $x, y$  and suppose  $x > y$ . Then  $x - y = 100$  and we need to choose  $x, y$  in such a way that  $xy$  is as small as possible. If we replace  $x$  with  $100 + y$ , then this means we need to minimize

$$f(y) = 100y + y^2.$$

The critical point of  $f(y)$  is  $y = -50$ , and since  $f$  is a parabola we know that this must be the location of the absolute minimum value on the graph. So the numbers we need are  $50, -50$ .

**Example 4.4.2 Minimizing material costs.** A box with an open top and a square base is to be built using material that costs \$5 per square meter. The box must have a volume of  $32\,000\text{ cm}^3$ . Find the dimensions of the box that will be cheapest to build to these specifications.

**Solution.** There are two things we need to do here:

1. Find the quantity to be optimize.
2. Solve the resulting optimization problem.

We need to minimize costs, which means we need to minimize how much material is used to construct this box. Hence we must minimize the surface area of the box. If we let  $x$  denote the length of the base of the box and  $h$  the height of the box, then the surface area is given by

$$x^2 + 4xh.$$

Now, we can't minimize this yet because we have too many variables. However, we can solve for  $h$  in terms of  $x$  by using the volume constraint. Since the volume of the box is  $x^2h = 32000$ , this gives  $h = \frac{32000}{x^2}$ . So the quantity we need to minimize is

$$f(x) = x^2 + 4xh = x^2 + \frac{128000}{x}.$$

To minimize this, we find the critical points by solving  $f'(x) = 0$ . This gives, after some algebra,  $x = 40$ . Now we need to be careful here! We only have a critical point at this step. We need to make an argument now as to why this critical point should be an absolute minimum of  $f(x)$ . One way to do this is by setting up a sign chart for  $f'(x)$ . If we do this, then we see that  $f' < 0$  for  $x$  in  $(0, 40)$ , while  $f' > 0$  for  $x$  in  $(40, \infty)$ . If we think about what this tells us of the graph of  $f(x)$ , we see that  $x = 40$  must minimize  $f(x)$ , at least for  $x > 0$ . So the dimensions that minimize the cost of building this box are  $40 \times 40 \times 20$ .

**Example 4.4.3 Distance on an ellipse.** Find the point(s) on the ellipse  $4x^2 + y^2 = 4$  that are farthest from the point  $(1, 0)$ .

**Solution.** This is an optimization problem since we're trying to maximize distance. There are two things we need to do:

1. Find the function we need to optimize.
2. Find the extrema.

The function we need to optimize is just the distance function. In particular, if  $(x, y)$  is a point on the ellipse then we need to maximize

$$d = \sqrt{(x-1)^2 + (y-0)^2}.$$

Now, we have a constraint that  $(x, y)$  must satisfy; namely, this must lie on the ellipse  $4x^2 + y^2 = 4$ . This means that  $y^2 = 4 - 4x^2$ , and if we plug this into our distance function we get

$$d = \sqrt{(x-1)^2 + (y-0)^2} = \sqrt{(x-1)^2 + 4 - 4x^2}.$$

Now here's a trick we can use: maximizing  $d$  is the same thing as maximizing  $d^2$ , but  $d^2$  is *much* nicer to work with algebraically. So instead of maximizing  $d$ , we'll maximize the function

$$f(x) = d^2 = (x-1)^2 + 4 - 4x^2.$$

We'll start by finding it's local extrema, which means we need to find the critical points. These are the solutions of  $f'(x) = 0$ . Since

$$f'(x) = 2x - 2 - 8x = -6x - 2,$$

we see that the only critical point is  $x = -\frac{1}{3}$ . Since  $f''(x) = -6 < 0$ , this means that  $x = -\frac{1}{3}$  gives us a local maximum of  $f(x)$  by the second derivative test. In fact, we can go further: this must be an absolute maximum of  $f(x)$ , since  $f(x)$  is always concave down (it's actually a parabola).

So the point on the ellipse  $4x^2 + y^2 = 4$  farthest from  $(1, 0)$  has  $x$ -coordinate equal to  $-\frac{1}{3}$ . This means that the corresponding  $y$ -coordinate is  $y = \pm \frac{2\sqrt{5}}{3}$ , so the points on the ellipse that are the farthest from  $(1, 0)$  are the points  $(-\frac{1}{3}, -\frac{2\sqrt{5}}{3}), (-\frac{1}{3}, \frac{2\sqrt{5}}{3})$ .

## 4.5 Newton's Method

Consider a differentiable function  $f(x)$ , and suppose we want to find a root of  $f(x)$ , which is a number  $x_0$  such that  $f(x_0) = 0$ . In some cases this is very easy (like  $f(x) = x^2 + 2x + 1$ ), but in others this may be more complicated (such as  $f(x) = x^5 - x - 1$ ). So we'd like to find a way to *approximate* the root  $x_0$ .

**Newton's Method** begins as follows: pick some starting point, or guess,  $x_1$ . Then draw the tangent line  $L$  to the graph of  $f(x)$  at  $x_1$ . Now, the  $x$ -intercept of this line may not be a root, but hopefully it'll be closer to the root. So call this point  $x_2$ . We can solve for  $x_2$  (if  $f'(x_1) \neq 0$ ) to get

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

Now, at this point  $x_2$  is (hopefully!) a better approximation to the root  $x_0$  than  $x_1$ , and we can run through the same procedure to get a new point  $x_3$ .

In general, we can get a **sequence** of approximations  $x_1, x_2, \dots, x_n$  to the root  $x_0$  by using the formula

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},$$

as long as  $f'(x_n) \neq 0$ .

**Example 4.5.1 Approximating a root.** Find the third approximation given by Newton's Method for the root of  $f(x) = x^5 - x - 1$ , using  $x_1 = -1$ .

**Solution.** First, note that  $f'(x) = 5x^4 - 1$ . Then

$$\begin{aligned} x_2 &= x_1 - \frac{f(x_1)}{f'(x_1)} \\ &= -1 - \frac{-1}{4} \\ &= 3, \end{aligned}$$

so  $x_2 = 3$ . Now we run through this method again to get  $x_3$ :

$$\begin{aligned} x_3 &= x_2 - \frac{f(x_2)}{f'(x_2)} \\ &= 3 - \frac{239}{404} \\ &= \frac{973}{404}. \end{aligned}$$

So  $x_3 = \frac{973}{404}$ .

**Example 4.5.2 Approximating  $\pi$ .** Approximate  $\pi$  using Newton's method.

**Solution.** We need to find a function  $f(x)$  for which  $f(\pi) = 0$ . Perhaps the easiest is  $f(x) = \sin x$ . Now, to use Newton's method we also need a starting guess. We'll pick  $x_1 = 3$  since this is close to  $\pi$ . Then

$$x_2 = x_1 - \frac{\sin x_1}{\cos x_1} = 3 - \tan 3$$

so  $x_2 \approx 3.142546543$ . Similarly,

$$\begin{aligned} x_3 &= x_2 - \tan x_2 \approx 3.141592653 \\ x_4 &\approx 3.141592654 \end{aligned}$$

Since these approximations are so close, we estimate that  $\pi \approx 3.14159265$ .



**Example 4.5.3 A Babylonian problem.** Use Newton's method to find an algorithm for computing  $\sqrt{a}$ .

**Solution.** To use Newton's method, we need to come up with a function  $f(x)$  whose root is  $\sqrt{a}$ . A simple choice for this is  $f(x) = x^2 - a$ , since  $f(\sqrt{a}) = 0$ . Now, if we're getting some sequence  $x_1, \dots, x_n$  from Newton's method then the next term in the sequence is

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - a}{2x_n}.$$

After some algebra, we can rewrite this as

$$x_{n+1} = \frac{1}{2} \left( x_n + \frac{a}{x_n} \right).$$

We can actually test this algorithm out. Say we want to approximate  $\sqrt{5}$ . A reasonable guess would be  $x_1 = 2$ , since this should be close to  $\sqrt{5}$ . Then we get

$$\begin{aligned} x_2 &= \frac{1}{2} \left( 2 + \frac{5}{2} \right) = 2.25 \\ x_3 &= \frac{1}{2} \left( 2.25 + \frac{5}{2.25} \right) \approx 2.236111111 \\ x_4 &\approx 2.236067978 \\ x_5 &\approx 2.236067977 \end{aligned}$$

So it looks like  $\sqrt{5} \approx 2.23606797$ , and indeed this is quite close to the actual value of  $\sqrt{5} = 2.236067977\dots$

We can go one step further for this algorithm. If  $x_1, x_2, \dots, x_n, \dots$  is the sequence we obtain using Newton's method for approximating  $\sqrt{a}$ , then as we saw earlier

$$x_{n+1} = \frac{1}{2} \left( x_n + \frac{a}{x_n} \right).$$

Now, assume that  $\lim_{n \rightarrow \infty} x_n$  exists and is nonzero, and call it  $s$ . Then

$$\lim_{n \rightarrow \infty} x_{n+1} = \frac{1}{2} \left( \lim_{n \rightarrow \infty} x_n + \frac{a}{\lim_{n \rightarrow \infty} x_n} \right)$$

becomes

$$s = \frac{1}{2} \left( s + \frac{a}{s} \right)$$

which we can rearrange to get

$$s^2 = \frac{1}{2}(s^2 + a),$$

or just  $s^2 = a$ . Hence  $s = \pm\sqrt{a}$ .

## 4.6 Antiderivatives

Suppose we're tracking a moving object, and through experimentation we know its acceleration to be equal to  $3 \frac{\text{m}}{\text{s}^2}$  to the right. Say we want to find the velocity of the mass at time  $t$ . Then if we set  $v(t)$  equal to the velocity and  $a(t)$  equal to the acceleration, we know that

$$v'(t) = a(t) = 3.$$

Therefore  $v(t)$  is a function whose derivative is 3. So one possible choice for  $v(t)$  is  $v(t) = 3t$ . This leads us to the definition of an **antiderivative**.

**Definition 4.6.1 Antiderivatives.** A function  $F(x)$  is called an antiderivative of the function  $f(x)$  if  $F'(x) = f(x)$  for all  $x$  in an interval  $I$ .

**Example 4.6.2 Multiple antiderivatives.** Find three different antiderivatives for  $f(x) = \cos x$ .

**Solution.** An antiderivative of  $\cos x$  is any function whose derivative is  $\cos x$ . So three antiderivatives could be

$$F_1(x) = \sin x, F_2(x) = \sin x + 4, F_3(x) = \sin x - 100.$$

From [Example 4.6.2](#) we see that any function of the form  $\sin x + C$  is an antiderivative of  $\cos x$ . In fact, this describes all possible antiderivatives of  $\cos x$ . This suggests the following theorem.

**Theorem 4.6.3 General Antiderivatives.** Let  $f(x)$  be some function with antiderivative  $F(x)$  on some interval  $I$ . Then the most general antiderivative of  $f(x)$  is

$$F(x) + C$$

where  $C$  is an arbitrary constant.

*Proof.* It might not seem like this statement requires a proof, but it does! We can check very easily that every function of the form  $F(x) + C$  is an antiderivative of  $f(x)$ , but how do we know every antiderivative takes this form? To prove this, let  $G(x)$  be an arbitrary antiderivative of  $f(x)$ . Then we need to show that  $G(x) = F(x) + C$ , or equivalently that  $G(x) - F(x) = C$ . We can do this by taking the derivative of  $G(x) - F(x)$  to get

$$G'(x) - F'(x) = f(x) - f(x) = 0,$$

and so by [Theorem 4.2.5](#) we know that  $G(x) - F(x)$  must be constant on  $I$ . Hence  $G(x) - F(x) = C$  for some constant  $C$ , or equivalently  $G(x) = F(x) + C$ . Therefore *every* antiderivative of  $f(x)$  can be written as  $F(x) + C$ .  $\square$

It might seem superfluous, but typically when dealing with models requiring us to find antiderivatives we'll want to find the most general antiderivative somewhere along the way. Moral of the story: *don't forget to add  $C$ .*

**Example 4.6.4 Finding antiderivatives.** Find the most general antiderivatives of the following functions:

1.  $x^3$ .
2.  $-\frac{1}{2}x^{\frac{5}{3}}$ .
3.  $\sqrt{x} - 3x^3 + x^{-2}$ .
4.  $\sin x + \cos(5x) + e^{4x}$ .

**Solution.** We have the following:

1. The general antiderivative of  $x^3$  is  $\frac{1}{4}x^4 + C$ .
2. The general antiderivative of  $-\frac{1}{2}x^{\frac{5}{3}}$  is  $-\frac{1}{2} \cdot \frac{3}{8}x^{\frac{8}{3}} + C$ .
3. The general antiderivative of  $\sqrt{x} - 3x^3 + x^{-2}$  is  $\frac{2}{3}x^{\frac{3}{2}} - \frac{3}{4}x^4 - \frac{1}{x} + C$ .
4. The general antiderivative of  $\sin x + \cos(5x) + e^{4x}$  is  $-\cos x + \frac{1}{5}\sin(5x) + \frac{1}{4}e^{4x} + C$ .

**Example 4.6.5 The antiderivative of  $\frac{1}{x}$ .** Find the general antiderivative of  $f(x) = \frac{1}{x}$ .

**Solution.** We have two cases we need to consider, since the domain of  $f(x)$  consists of two intervals. First, suppose that  $x$  is in  $(0, \infty)$ . Then we know that  $\frac{d}{dx} \ln x = \frac{1}{x}$ , so on this interval the most general antiderivative of  $\frac{1}{x}$  is  $\ln x + C_1$ .

Now let  $x$  be in  $(-\infty, 0)$ . Then  $\ln x$  isn't defined here. However, we can write

$$\frac{1}{x} = -\frac{1}{-x},$$

and if  $x < 0$  then  $-x > 0$  and  $\ln(-x)$  is defined, and in fact

$$\frac{d}{dx} \ln(-x) = -\frac{1}{(-x)} = \frac{1}{x}.$$

So the most general antiderivative of  $\frac{1}{x}$  on  $(-\infty, 0)$  is  $\ln(-x) + C_2$ . Putting this all together, we can say that the most general antiderivative of  $f(x) = \frac{1}{x}$  is given by

$$F(x) = \begin{cases} \ln x + C_1 & x > 0 \\ \ln(-x) + C_2 & x < 0 \end{cases}.$$

If we know that we're only selecting values of  $x$  with the same sign, then we can just say that the antiderivative of  $\frac{1}{x}$  is  $\ln|x| + C$ .

**Example 4.6.6 Antiderivatives involving secant.** Find the general antiderivatives of  $\sec^2 t$  and  $-2 \sec 3t \tan 3t$ .

**Solution.** The first antiderivative isn't too hard; it's just  $\tan t + C$ . The second is a little more complicated but still not too bad. Since  $\frac{d}{dt} \sec t = \sec t \tan t$ , a good guess for an antiderivative of  $-2 \sec 3t \tan 3t$  would be  $-2 \sec 3t$ . However, the derivative of this is  $-6 \sec 3t \tan 3t$ , so we're off by a factor of 3. So we just need to divide our guess by 3 to correct for this. Hence the general antiderivative of  $-2 \sec 3t \tan 3t$  is  $-\frac{2}{3} \sec 3t$ .

**Theorem 4.6.7 Antidifferentiation Formulas.** Let  $F(x)$  and  $G(x)$  denote antiderivatives of  $f(x)$  and  $g(x)$ , and let  $c$  be a constant. Then we have the following:

Function	Antiderivative
$cf(x)$	$cF(x)$
$f(x) + g(x)$	$F(x) + G(x)$
$x^n, n \neq -1$	$\frac{x^{n+1}}{n+1}$
$\frac{1}{x}$	$\ln x $
$e^{cx}, c \neq 0$	$\frac{1}{c}e^{cx}$
$\sin cx, \cos cx, c \neq 0$	$-\frac{1}{c} \cos cx, \frac{1}{c} \sin cx$

**Table 4.6.8:** Particular antiderivatives of functions

*These are by no means the only antiderivatives that you will need to deal with, but they are probably the most common.*

**Example 4.6.9 Falling objects.** An object is dropped from a height of 100 m above sea level and falls with downward acceleration equal to  $9.8 \frac{\text{m}}{\text{s}^2}$ . Find the height  $h(t)$  of the object  $t$  seconds after it's dropped.

**Solution.** Take downward to be the negative direction and sea level to be  $h = 0$ . Since acceleration is the second derivative of position, then the position (i.e. height) of the object should be the second *antiderivative* of acceleration. At this step, it's tempting to say that since  $h''(t) = -9.8$ , we have

$$h'(t) = -9.8t \quad \text{and} \quad h(t) = -4.9t^2.$$

And indeed,  $-4.9t^2$  is a function whose second derivative is  $-9.8$ . However, we have a slight problem here. We know that  $h(0) = 100$  and  $h'(0) = 0$ . If  $h(t) = -4.9t^2$ , then it's impossible for  $h(0)$  to equal 100. The problem here is we forgot about the arbitrary constants.

To get an accurate expression for  $h(t)$ , we go back to  $h''(t) = -9.8$ . Then  $h'(t) = -9.8t + C_1$  for some constant  $C_1$ . Since  $h'(0) = 0$ , this forces  $C_1 = 0$ . So we have  $h'(t) = -9.8t$ . Now we antidifferentiate one more time to get  $h(t) = -4.9t^2 + C_2$ . Since  $h(0) = 100$ , this forces  $C_2 = 100$ . So

$$h(t) = -4.9t^2 + 100.$$

Some functions that do not have an obvious antiderivative can be simplified through algebra into a form that is perhaps more helpful.

**Example 4.6.10 A tricky antiderivative.** Find the most general antiderivative of

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

**Solution.** It's tough to think of a function whose derivative is  $f(x)$ , though since the denominator is  $1 + x^2$  it seems likely that this antiderivative will involve  $\tan^{-1} x$  in some way. In order to actually find the antiderivative, we'll rewrite  $f(x)$  into a more convenient form. First, note that the numerator is very close to the denominator, which means we can almost cancel it out. So we'll split up the numerator as follows:

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

We can find the antiderivative of this term by term, so the most general antiderivative of  $f(x)$  is

$$F(x) = \tan^{-1} x + x + C.$$

At this point, it's natural to think about what an antiderivative means geometrically since the derivative of a function has such a nice interpretation. One approach to this is to think about **net change** of the antiderivative. Let  $F(x)$  be an antiderivative of  $f(x)$  on some interval containing  $a, b$ . The net change of  $F(x)$  from  $x = a$  to  $x = b$  is given by

$$F(b) - F(a).$$

Now,  $F'(x) = f(x)$  represents instantaneous rate of change, not net change. However, if we can somehow add these instantaneous changes from  $x = a$  to  $x = b$ , they should accumulate until they give the net change of  $F$  from  $a$  to  $b$ . Graphically, adding these instantaneous rates of change looks like it's giving us the *area under  $f(x)$* , which suggests the following geometric interpretation: *the area under  $f(x)$  from  $x = a$  to  $x = b$  is just the net change of an antiderivative  $F(x)$  from  $x = a$  to  $x = b$ .* Much of the next chapter is devoted to making this idea precise.

# Chapter 5

## Integrals

If a function is continuous, then we can always find its antiderivative via integration. The goal of this chapter is to introduce and refine this concept.

### 5.1 Areas Under Curves

We know how to find areas of simple shapes, such as rectangles, triangles and circles. But how can we find the area of a more complicated shape? Say, for example, the area under  $y = x^3$  from  $x = 0$  to  $x = 1$ ? Well, if we don't know it exactly then we can at least approximate it using a shape we do know. For example, using rectangles.

A simple way to divide the graph of  $y = x^3$  using rectangles is to pick several equally spaced points in the interval  $[0, 1]$ , say

$$x_0 = 0, x_1 = \frac{1}{3}, x_2 = \frac{2}{3}, \quad \text{and} \quad x_3 = 1,$$

and then draw rectangles using these points to determine the base of each rectangle and the height of the graph above the points to determine the height of each rectangle. In particular, we will have three rectangles; the first with base  $[0, \frac{1}{3}]$ , the second with base  $[\frac{1}{3}, \frac{2}{3}]$  and the third with base  $[\frac{2}{3}, 1]$ . Now, the height of each rectangle comes from a point on the graph above the corresponding interval. To make these specific (and relatively straightforward), we'll use **right endpoints** to determine the height of each rectangle. So the height of each respective rectangle is given by  $(\frac{1}{3})^3$ ,  $(\frac{2}{3})^3$  and by  $1^3$ .

Now, if we add up the areas of these rectangles we get an approximation for the area under the graph. Call these areas  $A_1, A_2, A_3$ . So the area under the graph of  $y = x^3$  from  $x = 0$  to  $x = 1$  is about

$$A_1 + A_2 + A_3 = \frac{4}{9}.$$

There's nothing stopping us from going further here. For example, we pick  $n$  rectangles, each with base length  $\frac{1}{n}$  and heights determined by right endpoints as above, then the area  $R_n$  of all of the rectangles will be an approximation (in fact, an overestimate) of the area of  $y = x^3$ . The bases of these rectangles are given by

$$[0, \frac{1}{n}], [\frac{1}{n}, \frac{2}{n}], \dots, [\frac{n-1}{n}, \frac{n}{n}].$$

So the areas of these rectangles are

$$\frac{1}{n} \left(\frac{1}{n}\right)^3, \frac{1}{n} \left(\frac{2}{n}\right)^3, \dots, \frac{1}{n} \left(\frac{n-1}{n}\right)^3, \frac{1}{n} \left(\frac{n}{n}\right)^3.$$

Adding these up, we get

$$R_n = \frac{1}{n} \frac{1^3 + 2^3 + 3^3 + \cdots + n^3}{n^3}$$

or just

$$R_n = \frac{1^3 + 2^3 + \cdots + n^3}{n^4}.$$

So if we sent  $n$  to  $\infty$ , we should get the exact area. In other words, the area under  $y = x^3$  from  $x = 0$  to  $x = 1$  is given by

$$\lim_{n \rightarrow \infty} R_n = \lim_{n \rightarrow \infty} \frac{1^3 + 2^3 + \cdots + n^3}{n^4}.$$

Actually finding this limit requires some trickery. It's not at all obvious, but we can write

$$1^3 + 2^3 + \cdots + n^3 = \left[ \frac{n(n+1)}{2} \right]^2.$$

Hence

$$\begin{aligned} \lim_{n \rightarrow \infty} R_n &= \lim_{n \rightarrow \infty} \frac{1}{n^4} \left[ \frac{n(n+1)}{2} \right]^2 \\ &= \lim_{n \rightarrow \infty} \frac{1}{4} \frac{n^2(n^2 + 2n + 1)}{n^4} \\ &= \lim_{n \rightarrow \infty} \frac{1}{4} \frac{n^4 + 2n^3 + n^2}{n^4} \\ &= \frac{1}{4}. \end{aligned}$$

So the exact area under  $y = x^3$  from  $x = 0$  to  $x = 1$  is  $\frac{1}{4}$ . Note that this matches up with our intuition from the last chapter: the area should be equal to the net change of an antiderivative. In this case, an antiderivative of  $x^3$  is  $\frac{1}{4}x^4$ , and the net change of  $\frac{1}{4}x^4$  from  $x = 0$  to  $x = 1$  is exactly  $\frac{1}{4}$ .

The process we used above for  $y = x^3$  can be carried out for any continuous function. If we have a function defined on the interval  $[a, b]$ , we can approximate the area under the graph by using rectangles. First, we split  $[a, b]$  into  $n$  different subintervals  $[x_0, x_1], \dots, [x_{n-1}, x_n]$ , each of width  $\Delta x = \frac{b-a}{n}$  and where

$$\begin{aligned} x_0 &= a \\ x_1 &= a + \Delta x = a + \frac{b-a}{n} \\ x_2 &= a + 2\Delta x = a + 2\frac{b-a}{n} \\ &\vdots \\ x_n &= a + n\Delta x = b \end{aligned}$$

The approximate area under the graph of  $y = f(x)$  from  $x = a$  to  $x = b$  is then

$$R_n = f(x_1)\Delta x + f(x_2)\Delta x + \cdots + f(x_n)\Delta x.$$

We can now define the area under a graph of a continuous function. Again, the idea is to approximate the graph with more and more rectangles.

**Definition 5.1.1 Area.** The **area** under the graph of a continuous function  $y = f(x)$  is the limit of the sum of the areas of the approximating rectangles:

$$A = \lim_{n \rightarrow \infty} R_n = \lim_{n \rightarrow \infty} [f(x_1)\Delta x + \cdots + f(x_n)\Delta x].$$

Although the previous definition used right endpoints to find the height of each rectangle, this doesn't have to be the case. We can use left endpoints, midpoints or any other **sample points**  $x_1^*, x_2^*, \dots, x_n^*$ , where each  $x_i^*$  is a point in the interval  $[x_{i-1}, x_i]$ .

Since finding areas in this way requires writing complicated sums, it seems like a good idea at this point to introduce notation for writing large summations. We call this **sigma notation**. Sigma notation takes the following form:

$$\sum_{i=\text{start}}^{\text{finish}} (\text{terms involving } i).$$

To evaluate such an expression, just plug in the different values for  $i$  from start to finish and add up what you get.

**Example 5.1.2 Evaluating a sum.** Determine  $\sum_{i=-2}^3 (4 - 3i)$ .

**Solution.** We start at  $i = -2$  and finish at  $i = 3$ . Plugging all of these values in for  $i$ , we get the following table:

$i$	$4 - 3i$
-2	10
-1	7
0	4
1	1
2	-2
3	-5

So we can say that

$$\begin{aligned} \sum_{i=-2}^3 (4 - 3i) &= 10 + 7 + 4 + 1 - 2 - 5 \\ &= 15. \end{aligned}$$

**Example 5.1.3 Using sigma notation.** Use sigma notation to write the sum of all squares from 31 to 245.

**Solution.** We need to start at  $i = 31$  and stop at  $i = 245$ . So the first component of this sum using sigma notation looks like  $\sum_{i=31}^{245}$ . Now we just need to figure out what goes inside of it. Since we're adding squares of numbers, we'll just put  $i^2$  inside of this sum. Hence the sum of all squares from 31 to 245 can be denoted

$$\sum_{i=31}^{245} i^2.$$

Sigma notation has some useful properties, since it's just another way to write sums.

**Theorem 5.1.4 Properties of Sigma Notation.** Let  $X = \sum_{i=a}^b x_i$  and let  $Y = \sum_{i=a}^b y_i$ . Let  $c$  be a constant. Then the following are true:

$$\sum_{i=a}^b c = c + \cdots + c = (b - a + 1)c$$

$$\begin{aligned}
\sum_{i=a}^b cx_i &= c \sum_{i=a}^b x_i = cX \\
\sum_{i=a}^b (x_i + y_i) &= \sum_{i=a}^b x_i + \sum_{i=a}^b y_i = X + Y \\
\sum_{i=a}^b (x_i - y_i) &= \sum_{i=a}^b x_i - \sum_{i=a}^b y_i = X - Y \\
\sum_{i=a}^k a_i + \sum_{i=k+1}^b a_i &= \sum_{i=a}^b a_i.
\end{aligned}$$

In other words, we can break sigma notation up over addition and subtraction, and we can move constants outside of it.

**Example 5.1.5 Areas using sigma notation.** Write down a limit that gives the area under  $f(x) = \sin x$  from  $x = 1$  to  $x = 3$  using sigma notation.

**Solution.** First, note that we can rewrite [Definition 5.1.1](#) using sigma notation as

$$A = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i) \Delta x.$$

We know what  $f(x)$  is, so we just need to identify  $x_i$  and  $\Delta x$ .  $\Delta x$  isn't too bad here, it's just

$$\Delta x = \frac{3-1}{n} = \frac{2}{n}.$$

The  $x_i$  are a little more complicated, but we'll still use right endpoints to find them. So we have

$$\begin{aligned}
x_0 &= 1 \\
x_1 &= 1 + \frac{2}{n} = \frac{n+2}{n} \\
x_2 &= 1 + 2\frac{2}{n} = \frac{n+4}{n} \\
x_3 &= 1 + 3\frac{2}{n} = \frac{n+6}{n} \\
&\vdots \\
x_n &= 3
\end{aligned}$$

In general,

$$x_i = 1 + i\frac{2}{n} = \frac{n+2i}{n}$$

for  $1 \leq i \leq n$ .

Now we have everything we need to write our limit. Hence the area under  $\sin x$  from 1 to 3 is given by

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \sin\left(\frac{n+2i}{n}\right) \frac{2}{n}.$$

Just as the problem of finding the slope of a tangent line can be viewed as a rate of change problem, the problem of finding areas can be viewed as an *accumulation problem*.



**Example 5.1.6 Finding work done.** In physics, the work done by a constant force  $F$  on some particle that moves a displacement  $d$  is given by the formula

$$W = Fd.$$

How should we define the work done by a variable force?

**Solution.** The main idea here is to treat the variable force as the combination of several constant forces. To be specific, let  $F(x)$  denote our force (which depends on position  $x$ ) and suppose it acts on a particle moving from  $x = a$  to  $x = b$ .  $F(x)$  itself may not be constant, but if we break  $[a, b]$  up into  $n$  subintervals of the form  $[x_{i-1}, x_i]$  each of length  $\Delta x = \frac{b-a}{n}$ , then  $F(x)$  should be nearly constant on each subinterval. So we can pick points  $x_i^*$  from each subinterval and say that  $F(x) \approx F(x_i^*)$  for all  $x$  in  $[x_{i-1}, x_i]$ .

Now we can approximate the work done by  $F$  as the particle moves through the subinterval  $[x_{i-1}, x_i]$ . The work done should be about equal to

$$F(x_i^*)\Delta x.$$

Now, if we add up these estimates for each subinterval, we see that the total work done should be approximated by

$$\sum_{i=1}^n F(x_i^*)\Delta x,$$

which suggests that we should define the exact work done to be

$$W = \lim_{n \rightarrow \infty} \sum_{i=1}^n F(x_i^*)\Delta x.$$

Note that [Example 5.1.6](#) was not written in terms of finding areas, but still involved the limit from [Definition 5.1.1](#). This tells us that the ideas in this section have a wider use than in just finding areas under curves.

## 5.2 The Definite Integral

The process used in [Section 5.1](#) can be generalized quite a bit. Start with a function  $f(x)$  defined on some interval  $[a, b]$ , and divide  $[a, b]$  into  $n$  subintervals

$$[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$$

where  $x_0 = a, x_n = b$  and each subinterval has length  $\Delta x_i$  (i.e.

$$\Delta x_i = x_i - x_{i-1}$$

). This is known as a **partition** of  $[a, b]$ . Then we can choose sample points  $x_1^*$  from  $[x_0, x_1]$ ,  $x_2^*$  from  $[x_1, x_2]$ , all the way up to  $x_n^*$  from  $[x_{n-1}, x_n]$ . Note that the only restriction we're placing on the sample point  $x_i^*$  is that it must come from the interval  $[x_{i-1}, x_i]$ , but it may be *any* point in this interval.

Using all of this, we can define the **Riemann sum** associated to  $f$  on this partition:

$$\sum_{i=1}^n f(x_i^*)\Delta x_i.$$

Although this can be more complicated than our definition in [Definition 5.1.1](#), it still approximates the area under  $f(x)$  from  $x = a$  to  $x = b$ . And as each subinterval in this partition gets smaller (i.e. as  $\Delta x_i \rightarrow 0$  or equivalently as  $n \rightarrow \infty$ ), then this approximation should become exact, at least if  $f(x)$  is nice enough.

**Definition 5.2.1 The Definite Integral.** Let  $f(x)$  be a function defined on  $[a, b]$ . Then the **definite integral of  $f(x)$  from  $a$  to  $b$**  is the number

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i^*) \Delta x_i,$$

provided that the limit exists (and that  $\Delta x_i \rightarrow 0$ ). If this limit exists, we say that  $f(x)$  is **integrable** on  $[a, b]$ .

We can use [Definition 5.2.1](#) to rewrite our area definition: the area under  $y = f(x)$  from  $x = a$  to  $x = b$  is defined to be the definite integral of  $f$  on  $[a, b]$ , assuming it exists. For a function to be integrable, the particular partition of  $[a, b]$  that we use *cannot affect the limit* in [Definition 5.2.1](#). This gives us a great deal of freedom in approximating the definite integral/area under a curve: when choosing sample points, we can use left endpoints; right endpoints; midpoints and more. However, if  $f$  is integrable then we lose nothing by just using right endpoints and subintervals of equal length, as in [Definition 5.1.1](#).

**Theorem 5.2.2 Integrability of Piecewise Continuous Functions.** *Let  $f(x)$  be a piecewise function on  $[a, b]$  with finitely many jump discontinuities. Then  $f$  is integrable on  $[a, b]$ .*

**Example 5.2.3 Expressing work done.** A particle moving from  $x = -1$  to  $x = 3$  is acted upon by the force  $F(x) = 3x$ . Find the work done by the force on this particle.

**Solution.** Using [Example 5.1.6](#), we see that the work done should be

$$W = \lim_{n \rightarrow \infty} \sum_{i=1}^n 3x_i^* \cdot \frac{4}{n}$$

which we can rewrite as

$$W = \int_{-1}^3 3x dx.$$

So work is a definite integral.

In this case, we can also find the work done. Since the definite integral represents area, and the definite integral of  $3x$  exists (since this is continuous), then we can say that

$$W = \int_{-1}^3 3x dx = \frac{1}{2}(1)(-3) + \frac{1}{2}(3)(9) = 12.$$

**Example 5.2.4 A radical integral.** Determine  $\int_0^2 \sqrt{4-x^2} dx$ .

**Solution.** We won't be able to solve this integral algebraically until Calculus 2, but we can solve it geometrically instead. If we graph  $y = \sqrt{4-x^2}$ , we get the upper half of the circle of radius 2 centered at the origin. So the integral given must be exactly one-fourth the area of this circle, so

$$\int_0^2 \sqrt{4-x^2} dx = \frac{1}{4}\pi(2^2) = \pi.$$

Since the definite integral is defined as a limit of sums, it shares many properties with sums. In particular, we have the following version of [Theorem 5.1.4](#)

**Theorem 5.2.5 Properties of the Definite Integral.** Let  $f(x)$  and  $g(x)$  be integrable on  $[a, b]$  and let  $c$  be a constant. Let  $t$  be any number in  $[a, b]$ . Then the following are true:

$$\begin{aligned}\int_a^b c \, dx &= c(b - a) \\ \int_a^b cf(x) \, dx &= c \int_a^b f(x) \, dx \\ \int_a^b (f(x) + g(x)) \, dx &= \int_a^b f(x) \, dx + \int_a^b g(x) \, dx \\ \int_a^b (f(x) - g(x)) \, dx &= \int_a^b f(x) \, dx - \int_a^b g(x) \, dx \\ \int_a^t f(x) \, dx + \int_t^b f(x) \, dx &= \int_a^b f(x) \, dx.\end{aligned}$$

In other words, we can break definite integrals up over addition and subtraction, and we can move constants outside of it.

Computing integrals in practice often involves setting up and approximating Riemann sums, i.e. sums of the form  $\sum_{i=1}^n f(x_i^*)\Delta x_i$ . These sums depend on how we choose the sample points  $x_i^*$  from each interval  $[x_{i-1}, x_i]$ . We've mostly worked with  $x_i^* = x_i$ , or right endpoints. We could also choose  $x_i^* = x_{i-1}$ , using left endpoints. However, it's often more accurate to use *midpoints*:  $x_i^* = \frac{x_{i-1} + x_i}{2}$ . The reason for this is that if we choose midpoints of intervals as our sample points, these tend to be less affected than either endpoint if the function is increasing or decreasing. In particular, we have the **midpoint rule**.

**Definition 5.2.6 Midpoint Rule.** Let  $f(x)$  be a continuous function. Then

$$\int_a^b f(x) \, dx \approx M_n = \sum_{i=1}^n f(x_i^*)\Delta x$$

where

$$\begin{aligned}x_i^* &= \frac{x_{i-1} + x_i}{2} \\ \Delta x &= \frac{b - a}{n}.\end{aligned}$$

Note that this rule calls for evenly dividing  $[a, b]$  into a partition  $[x_{i-1}, x_i]$  where each subinterval has equal length  $\Delta x$ .

**Example 5.2.7 Approximating an integral.** Approximate  $\int_1^3 \frac{1}{x} \, dx$  using Riemann sums with  $n = 4$ .

**Solution.** For this integral, we have

$$\begin{aligned}\Delta n &= \frac{3 - 1}{4} = \frac{1}{2} \\ x_0 &= 1 \\ x_1 &= \frac{3}{2} \\ x_2 &= 2 \\ x_3 &= \frac{5}{2}\end{aligned}$$

$$x_4 = 3$$

We will have different estimates using left endpoints, right endpoints and midpoints. Using left endpoints, we have

$$\int_1^3 \frac{1}{x} dx \approx L_4 = \frac{1}{2} \left( 1 + \frac{2}{3} + \frac{1}{2} + \frac{2}{5} \right) = 1.283\dots$$

Using midpoints, we have

$$\int_1^3 \frac{1}{x} dx \approx M_4 = \frac{1}{2} \left( \frac{4}{5} + \frac{4}{7} + \frac{4}{9} + \frac{4}{11} \right) \approx 1.09.$$

Using right endpoints, we have

$$\int_1^3 \frac{1}{x} dx \approx R_4 = \frac{1}{2} \left( \frac{2}{3} + \frac{1}{2} + \frac{2}{5} + \frac{1}{3} \right) = 0.95.$$

Geometrically, we see that

$$0.95 = R_4 < \int_1^3 \frac{1}{x} dx < L_4 = 1.283\dots$$

The exact value of this integral is actually  $\ln 3$ , and we see that  $M_4$ , the midpoint estimate, is clearly the best approximation to the integral in this case.

**Theorem 5.2.8 More Properties of the Definite Integral.** *Let  $f(x), g(x)$  be integrable functions on  $[a, b]$ . Then*

1.  $\int_a^b f(x) dx = -\int_b^a f(x) dx$
2.  $\int_a^a f(x) dx = 0$
3. If  $f(x) \geq g(x)$  on  $[a, b]$ , then  $\int_a^b f(x) dx \geq \int_a^b g(x) dx$ .
4. If  $m \leq f(x) \leq M$  on  $[a, b]$ , then  $m(b-a) \leq \int_a^b f(x) dx \leq M(b-a)$ .

## 5.3 Evaluating Integrals

Just as we did with derivatives, we want to find a way to calculate integrals without having to compute limits every time.

**Theorem 5.3.1 Evaluation Theorem.** *Let  $f(x)$  be continuous on  $[a, b]$  and let  $F(x)$  be a single antiderivative of  $f(x)$  on this interval. Then*

$$\int_a^b f(x) dx = F(b) - F(a).$$

*Proof.* First, partition  $[a, b]$  into  $[x_0, x_1], \dots, [x_{n-1}, x_n]$ . Then we can say that

$$\begin{aligned} F(b) - F(a) &= F(x_n) - F(x_0) \\ &= F(x_n) - F(x_{n-1}) + F(x_{n-1}) - F(x_{n-2}) + \dots + F(x_1) - F(x_0) \\ &= \sum_{i=1}^n [F(x_i) - F(x_{i-1})]. \end{aligned}$$

By the Mean Value Theorem, we can rewrite each term in the sum using  $F'$ :

$$F(x_i) - F(x_{i-1}) = F'(x_i^*)(x_i - x_{i-1})$$

where  $x_i^*$  is just some point in  $[x_{i-1}, x_i]$ . So we have

$$\begin{aligned} F(b) - F(a) &= \sum_{i=1}^n F'(x_i^*)(x_i - x_{i-1}) \\ &= \sum_{i=1}^n f(x_i^*)\Delta x. \end{aligned}$$

If we now send  $n \rightarrow \infty$  (i.e.,  $\Delta x \rightarrow 0$ ), then the above becomes

$$F(b) - F(a) = \int_a^b f(x) dx. \quad \square$$

**Theorem 5.3.1** tells us that if we want to evaluate the definite integral of a continuous function, we need to first find an antiderivative. We can also think of this result as follows: the definite integral of a rate of change is just the net change of the antiderivative.

**Example 5.3.2 Modeling a population.** A honeybee population starts with 100 members and grows at a rate of  $n'(t)$  per week. Find the total population after 15 weeks in terms of  $n'(t)$ .

**Solution.** The total population after 15 weeks should be given by

$$\begin{aligned} (\text{total population}) &= (\text{initial population}) + (\text{net change}) \\ &= 100 + \int_0^{15} n'(t) dt \end{aligned}$$

As a useful shorthand, we introduce the notation  $F(b) - F(a) = F(x)]_a^b$ . So if  $F$  is an antiderivative of  $f$ , we can say that

$$\int_a^b f(x) dx = F(x)]_a^b.$$

**Example 5.3.3 Area under a rational function.** Find the area under  $f(t) = \frac{t^2-1}{t^4-1} dt$  from  $t = 0$  to  $t = \frac{1}{\sqrt{3}}$ .

**Solution.** The area is just the definite integral, which is

$$\begin{aligned} \int_0^{\frac{1}{\sqrt{3}}} f(t) dt &= \int_0^{\frac{1}{\sqrt{3}}} \frac{t^2-1}{t^4-1} dt \\ &= \int_0^{\frac{1}{\sqrt{3}}} \frac{1}{t^2+1} dt \\ &= \tan^{-1} t \Big|_0^{\frac{1}{\sqrt{3}}} \\ &= \tan^{-1} \left( \frac{1}{\sqrt{3}} \right) - \tan^{-1}(0) \\ &= \frac{\pi}{6}. \end{aligned}$$

The **average value** of a continuous function  $f(x)$  on an interval  $[a, b]$  is defined to be

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

**Example 5.3.4 Average force.** An object of mass 3 kg is acted upon by several forces that combine to give it an acceleration of  $a(t) = 4e^{-t} + \sin t$ ,

where  $a(t)$  is in units of  $\frac{\text{m}}{\text{s}^2}$  and  $t$  is in seconds. Find the average force acting on the mass from  $t = 1$  to  $t = 3$ .

**Solution.** We first need to find the net force acting on the particle. By Newton's Second Law, this is  $F_{\text{net}} = ma = 12e^{-t} + 4 \sin t$ . So the average force should be

$$\begin{aligned} \frac{1}{2} \int_1^3 (12e^{-t} + 4 \sin t) dt &= \frac{1}{2} [-12e^{-t} - 4 \cos t]_1^3 \\ &= \frac{-12e^{-3} - 4 \cos 3 + 12e^{-1} + 4 \cos 1}{2}. \end{aligned}$$

Because of the relationship between definite integrals and antiderivatives, we introduce a notation for antiderivatives that reinforces this connect.

**Definition 5.3.5 Indefinite Integral.** Let  $f(x)$  be a continuous function. The **indefinite integral** of  $f(x)$ , denoted  $\int f(x) dx$ , is the general antiderivative of  $f(x)$ . In symbols, if  $F(x)$  is an antiderivative of  $f(x)$  then

$$\int f(x) dx = F(x) + C.$$

**Example 5.3.6 Finding an indefinite integral.** Given that

$$\frac{d}{dx}[xe^x - e^x] = xe^x,$$

find  $\int [xe^x + \sec^2 x] dx$ .

**Solution.** We can find this integral by splitting it up, since the antiderivative of a sum is just the sum of the antiderivatives. So

$$\begin{aligned} \int [xe^x + \sec^2 x] dx &= \int xe^x dx + \int \sec^2 x dx \\ &= xe^x - e^x + \tan x + C. \end{aligned}$$

Note that using [Theorem 5.3.1](#) to find definite integrals relies on having a particular antiderivative of a function  $f(x)$ , but we've said nothing about whether or not such an antiderivative can exist. In the next section, we'll see how to use integrals to construct antiderivatives.

## 5.4 The Fundamental Theorem of Calculus

The previous section used antiderivatives to find definite integrals. Now we'll go the other direction and use integrals to define antiderivatives. We can do this as follows.

Let  $f$  be some function continuous on  $[a, b]$ . Then we can say that  $f$  must be continuous on every interval of the form  $[a, x]$  for  $x$  in  $[a, b]$ . This means that the definite integrals  $\int_a^x f(t) dt$  exist, and so we can use these to define a function, say  $g(x)$ .

**Example 5.4.1 Functions from integrals.** Let  $f(x) = 2 - x^2$  and define  $g(x) = \int_1^x f(t) dt$ . Find  $g(x)$ . Where is  $g(x)$  the largest?

**Solution.** First, note that we can compute  $\int_1^x f(t) dt$  by finding an antiderivative of  $f(t) = 2 - t^2$ . So

$$g(x) = \int_1^x (2 - t^2) dt = \left[ 2t - \frac{1}{3}t^3 \right]_{t=1}^x$$

or just

$$g(x) = 2x - \frac{1}{3}x^3 - \frac{5}{3}.$$

We can find where  $g$  is largest by using a derivative test or just geometry. Since  $g$  represents the area under  $f$  from 1 to  $x$ , it follows that this area will be largest at  $x = \sqrt{2}$ , since this is where  $f$  crosses the axis.

There's one important thing to take away from the previous example. The function  $g(x) = \int_1^x f(t) dt$  turned out to be an antiderivative of  $f(x)$ , so in particular  $g'(x) = f(x)$ . This is not a coincidence.

**Theorem 5.4.2 Fundamental Theorem of Calculus.** *Let  $f(x)$  be a continuous function defined on  $[a, b]$ .*

1. *Define  $g(x) = \int_a^x f(t) dt$  for  $x$  in  $[a, x]$ . Then  $g'(x) = f(x)$ , i.e.,  $g$  is an antiderivative of  $f$ .*

2. *Let  $F$  be a particular antiderivative of  $f$ . Then  $\int_a^b f(x) dx = F(b) - F(a)$ .*

*Proof.* We'll sketch the proof of the first part. First, assume that  $g(x) = \int_a^x f(t) dt$ . Then

$$\frac{g(x+h) - g(x)}{h} = \frac{1}{h} \int_x^{x+h} f(t) dt.$$

Now, if  $h$  is small and if  $x \leq t \leq x+h$ , then  $\int_x^{x+h} f(t) dt$  should be approximately equal to  $f(x)h$  since  $f$  is continuous. Therefore

$$\lim_{h \rightarrow 0} \frac{g(x+h) - g(x)}{h} = \lim_{h \rightarrow 0} \frac{f(x)h}{h} = f(x),$$

and so  $g'(x) = f(x)$ . □

The second part of [Theorem 5.4.2](#) is something we've seen before in [Theorem 5.3.1](#), but the first part is new. Essentially, this result says that integration and differentiation are *inverses* of each other.

**Example 5.4.3 Error function.** The **error function**, defined by

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt,$$

is one of the most important functions in statistics and other applications of mathematics (such as modeling the spread of a virus or the spread of heat). Find the derivative of  $\operatorname{erf}(x)$ .

**Solution.** By the Fundamental Theorem of Calculus, we have

$$\operatorname{erf}'(x) = \frac{2}{\sqrt{\pi}} e^{-x^2}.$$

Note that  $\operatorname{erf}'(x) \rightarrow 0$  as  $x \rightarrow \infty$ , which suggests that the error function "levels out" as  $x$  increases, which is in fact true.

**Example 5.4.4 Computing derivatives.** Let  $f(x) = \int_{\sin x}^1 \sqrt{t^2 + e^{-t}} dt$ . Find  $f'(x)$ .

**Solution.** We have

$$f'(x) = \frac{d}{dx} \int_{\sin x}^1 \sqrt{t^2 + e^{-t}} dt$$

$$\begin{aligned}
&= -\frac{d}{dx} \int_1^{\sin x} \sqrt{t^2 + e^{-t}} dt \\
&= -\sqrt{\sin^2 x + e^{-\sin x}} \cos x.
\end{aligned}$$

Here,  $\sin x$  acts like an inside function so we must use the chain rule in addition to the Fundamental Theorem of Calculus to find this derivative.

**Example 5.4.5 A business problem.** A company purchases equipment that depreciates at a rate  $f(t)$  and accumulates maintenance costs at a rate  $g(t)$ , where  $t$  denotes the number of months after purchase. The company wants to minimize the cost per month of the equipment due to depreciation and maintenance. After how much time should the company replace the equipment?

**Solution.** First, we need to figure out the cost per month due to depreciation and maintenance, call it  $C(t)$ . Then

$$C(t) = \frac{\text{total depreciation} - \text{total maintenance}}{t}.$$

Now, the total depreciation after  $t$  months is  $\int_0^t f(s) ds$ , and the total maintenance cost after  $t$  months is  $\int_0^t g(s) ds$ . So

$$C(t) = \frac{1}{t} \int_0^t [f(s) + g(s)] ds.$$

This is the function we need to minimize.

To minimize  $C(t)$ , we'll start by finding its critical points. Assuming that each of these functions is differentiable, we have

$$C'(t) = -\frac{1}{t^2} \int_0^t [f(s) + g(s)] ds + \frac{f(t) + g(t)}{t}$$

which reduces to

$$C'(t) = -\frac{1}{t} C(t) + \frac{f(t) + g(t)}{t}.$$

So the critical points occur at any value  $t = T$  for which  $C(T) = f(T) + g(T)$ , i.e., the cost per month due to depreciation and maintenance equals the rate of depreciation plus the rate of maintenance costs.

Now we need to see if we have a minimum or a maximum. To do so, we can use the second derivative test. After some algebra, we get that

$$C''(t) = 2 \frac{C(t) - (f(t) + g(t))}{t^2} + \frac{f'(t) + g'(t)}{t^2},$$

and so

$$C''(T) = \frac{f'(T) + g'(T)}{T^2} > 0.$$

Therefore  $T$  gives a local minimum. If we assume that  $C(t)$  is differentiable everywhere, then this local minimum should also be an absolute minimum since it's the only critical point.

## 5.5 Substitution

By [Theorem 5.4.2](#), we know that integration and differentiation are inverse operations. This means that differentiation rules, such as the chain rule and product rule, have corresponding *integration* rules. In this section we look at



the integration rule corresponding to the chain rule, known as  $u$ -substitution, and in Calculus 2 you'll see the integration rule corresponding to the product rule.

To see how the chain rule leads to an integration rule, suppose we have an expression  $F(g(x))$  where  $F, g$  are differentiable. Then the chain rule says that  $\frac{d}{dx}F(g(x)) = F'(g(x))g'(x)$ . Now, write  $F' = f$  and make the substitution  $u = g(x)$ . Then we get

$$f(g(x))g'(x) = f(u)\frac{du}{dx}.$$

Now we can integrate both sides of this equation with respect to  $x$  to get

$$\int f(g(x))g'(x) dx = \int f(u)\frac{du}{dx} dx = \int f(u) du.$$

This method is known as  **$u$ -substitution**.

**Theorem 5.5.1  $u$ -Substitution.** *Let  $u = g(x)$  be a differentiable function whose range is an interval  $I$ . Let  $f(u)$  be a continuous function on  $I$ . Then*

$$\int f(g(x))g'(x) dx = \int f(u) du.$$

Solving an integral  $\int f(x) dx$  by  $u$ -substitution typically proceeds as follows:

1. Choose some substitution  $u$  to simplify your integral. Your choice for  $u$  will most likely be an "inside function".
2. Rewrite your integral in terms of  $u$  and  $du$  instead of  $x$  and  $dx$ .
3. Solve the resulting integral in terms of  $u$  if possible, and then rewrite your answer in terms of the original variable  $x$ .

**Example 5.5.2 A simple substitution.** Determine  $\int e^{3x-2} dx$  by using the substitution  $u = 3x - 2$ .

**Solution.** We need to rewrite the integral in terms of  $u$ . An easy way to start is to replace  $3x - 2$  with  $u$ :

$$\int e^{3x-2} dx = \int e^u dx.$$

However, *we can't integrate yet*. This is because the integral needs to be written entirely in terms of  $u$  and  $du$ ; no terms involving  $x$ , including  $dx$ , can remain. to rewrite  $dx$  in terms of  $du$ , we'll find  $\frac{du}{dx}$  and then "solve" for  $dx$ :

$$\frac{du}{dx} = 3 \Rightarrow du = 3 dx \Rightarrow dx = \frac{du}{3}.$$

So our integral turns into

$$\begin{aligned} \int e^{3x-2} dx &= \int e^u \left( \frac{du}{3} \right) \\ &= \frac{1}{3} \int e^u du \\ &= \frac{1}{3} e^u + C. \end{aligned}$$

The last step is to replace  $u$  with  $3x - 2$ , and so

$$\int e^{3x-2} dx = \frac{1}{3} e^{3x-2} + C.$$

We can also check this by differentiating.

**Example 5.5.3 Simplifying an integral with substitution.** Determine  $\int (2 - 5t)^{-\frac{3}{2}} dt$ .

**Solution.** We can find  $\int u^{-\frac{3}{2}} du$  without too much trouble, but this integral is more complicated. However, we can use  $u$ -substitution to get our integral into this form. So we'll let  $u = 2 - 5t$ . Then  $du = -5 dt$ , which means  $dt = -\frac{du}{5}$  and so

$$\begin{aligned} \int (2 - 5t)^{-\frac{3}{2}} dt &= \int u^{-\frac{3}{2}} \left( -\frac{1}{5} du \right) \\ &= -\frac{1}{5} \int u^{-\frac{3}{2}} du \\ &= \frac{1}{5} \frac{2}{1} u^{-\frac{1}{2}} + C \\ &= \frac{2}{5} (2 - 5t)^{-\frac{1}{2}} + C. \end{aligned}$$

$u$ -substitution is often useful if one part of an integral can be written as the derivative of another part of the integral.

**Example 5.5.4 Integral of tangent.** Find the general antiderivative of  $\tan x$ .

**Solution.** We need to find  $\int \tan x dx$ . If we rewrite  $\tan x$  in terms of sine and cosine, we get

$$\int \tan x = \int \frac{\sin x}{\cos x} dx.$$

It's not obvious, but we can solve this using  $u$ -substitution since  $\sin x$  looks a lot like the derivative of  $\cos x$ . So we'll let  $u = \cos x$ , giving  $du = -\sin x dx$  or just  $dx = -\frac{du}{\sin x}$ . This simplifies our integral to

$$\int \tan x dx = \int -\frac{du}{u} = -\ln |u| + C = -\ln |\cos x| + C.$$

Since  $-\ln |\cos x| = \ln |\cos x|^{-1} = \ln |\sec x|$ , this shows that the general antiderivative of  $\tan x$  is

$$\int \tan x dx = \ln |\sec x| + C.$$

Note that  $u = \sin x$  *would not work* in this example, at least not without some nasty algebra.

We can also use  $u$ -substitution to find definite integrals.

**Example 5.5.5 Integrating with a natural log.** Determine  $\int_{e^2}^{e^4} \frac{dx}{x\sqrt{\ln x}}$ .

**Solution.** This integral looks awful, but the integrand contains  $\ln x$  and its derivative  $\frac{1}{x}$ , which suggests that the substitution  $u = \ln x$  might simplify things. So let  $u = \ln x$ . Then  $du = \frac{1}{x} dx$ , or  $dx = x du$ , and the integral becomes

$$\int_{e^2}^{e^4} \frac{dx}{x\sqrt{\ln x}} = \int_{x=e^2}^{x=e^4} \frac{x du}{x\sqrt{u}}$$

$$\begin{aligned}
&= \int_{x=e^2}^{x=e^4} \frac{du}{\sqrt{u}} \\
&= \left[ u \right]_{x=e^2}^{x=e^4} \\
&= \left[ \ln x \right]_{e^2}^{e^4} \\
&= 2.
\end{aligned}$$

In the last example we had to be careful about our limits of integration since they applied to  $x$  but not  $u$ . We can simplify the process a little by converting the  $x$  limits to  $u$  limits, as shown in the next example.

**Example 5.5.6 Area under  $x \sin x^2$ .** Find the area under the graph of  $y = x \sin x^2$  from  $x = 0$  to  $x = \sqrt{\pi}$ .

**Solution.** The area is given by  $\int_0^{\sqrt{\pi}} x \sin x^2 dx$ . We can solve this by  $u$ -substitution. Let  $u = x^2$ , our inside function. Then  $du = 2x dx$ , so  $dx = \frac{du}{2x}$  and our integral turns into

$$\int_{x=0}^{x=\sqrt{\pi}} \frac{1}{2} \sin u du.$$

Now we change the limits to be in terms of  $u$ , giving us  $u = 0$  and  $u = \pi$ . So the area is just

$$\begin{aligned}
\int_0^{\sqrt{\pi}} x \sin x^2 dx &= \int_0^{\pi} \frac{1}{2} \sin u du \\
&= \frac{1}{2} [-\cos u]_0^{\pi} \\
&= \frac{1}{2} (\cos 0 - \cos \pi) \\
&= 1.
\end{aligned}$$



## Chapter 6

# Techniques of Integration

This chapter focuses on several techniques for evaluating complicated integrals. As the tools in this chapter are in general more complicated than those that appeared in [Chapter 5](#), you should only use these if algebraic simplification or substitution are insufficient to compute an integral.

### 6.1 Integration by Parts

Suppose we want to integrate  $f(x) = x \sin x$ . This can't be simplified using algebra (since integrals don't split up over products!) and substitution probably won't be very helpful, since  $f(x)$  does not already contain an "inside function" that we could replace. The integration tools that we have are unable to compute this integral, so we need to develop more tools!

Recall that the substitution method for solving integrals from [Section 5.5](#) is derived from the chain rule. In fact, by [Theorem 5.4.2](#) every derivative rule available to us also provides a corresponding integration rule. So perhaps the integration rule corresponding to the product rule can help us to integrate products. We call this method *integration by parts*.

**Theorem 6.1.1 Integration by Parts.** *Let  $f(x)$  and  $g(x)$  be continuous functions, with  $g(x)$  differentiable. Then*

$$\int f(x)g'(x) dx = f(x)g(x) - \int g(x)f'(x) dx.$$

*Proof.* By the product rule, we have  $\frac{d}{dx}[f(x)g(x)] = f(x)g'(x) + f'(x)g(x)$ . Rearranging this, we have

$$f(x)g'(x) = [f(x)g(x)]' - f'(x)g(x).$$

Finally, we can integrate both sides to get

$$\int f(x)g'(x) dx = f(x)g(x) - \int g(x)f'(x) dx. \quad \square$$

A useful way to remember the integration by parts method is as follows:

$$\int u dv = uv - \int v du. \quad (6.1.1)$$

Integration by parts can be used to integrate products of functions *that can't already be evaluated by simpler means*.

**Example 6.1.2 Using the Integration by Parts Formula.** Compute  $\int x \sin x \, dx$ .

**Solution.** To use (6.1.1), we need to write  $\int x \sin x \, dx$  in the form  $\int u \, dv$ . Once we have  $u$  and  $dv$ , then we can set up the rest of the formula. One thing that can make this method tricky is the fact that we can have many different choices for  $u$  and  $dv$ ; the only thing we need for certain is for  $u \, dv = x \sin x$  here. However, if we look at the formula (6.1.1) then we'll see two new terms on the right hand side:  $du$  and  $v$ , together in the second integral. This means that we'll need to differentiate  $u$  and integrate  $v$ , and this will suggest what both should be.

Since we need to compute  $du$ , then we want  $u$  to be easy to differentiate. Similarly, we want  $dv$  to be easy to integrate. And since we'd like to be able to compute this integral, we'd also like  $v \, du$  to be simpler than  $u \, dv$ . Now return to  $x \sin x$ . Since  $x$  gets simpler if we differentiate it by  $\sin x$  doesn't, this suggests that we should set  $u = x$ . And since  $u = x$ , we must have  $dv = \sin x \, dx$  (since  $dv$  automatically becomes whatever is left over). With these choices, we have

$$\begin{aligned} du &= dx \\ v &= -\cos x \end{aligned}$$

Therefore

$$\begin{aligned} \int x \sin x \, dx &= \int u \, dv \\ &= uv - \int v \, du \\ &= -x \cos x - \int (-\cos x) \, dx \\ &= -x \cos x + \int \cos x \, dx \\ &= -x \cos x + \sin x + C. \end{aligned}$$

You may be wondering why we didn't include an arbitrary constant when we found  $v$  in Example 6.1.2. This is because including an arbitrary constant here doesn't affect the final answer. It's certainly not incorrect to include one here, but it adds more work than is already necessary.

When using the integration by parts method, it's usually a good idea to choose  $u$  to be the polynomial portion of your integrand if there is one. This is because polynomials become simpler when they're differentiated. However, there are examples where  $u$  should be something else.

**Example 6.1.3 Integrals Involving the Natural Logarithm.** Determine  $\int 3x^2 \ln x \, dx$ .

**Solution.** Since we need to integrate a product, and we can't easily do so with  $u$ -substitution or algebra, we'll try integration by parts. We could pick  $u = 3x^2$ , and this is certainly easy to differentiate so we can find  $du$ . However, doing so leaves us with  $dv = \ln x$ , which is tricky to integrate, but is easy to differentiate. So we'll try setting  $u = \ln x$  and  $dv = 3x^2 \, dx$ . Then

$$du = \frac{1}{x} \, dx$$

$$v = x^3$$

and so

$$\begin{aligned}\int 3x^2 \ln x \, dx &= uv - \int v \, du \\ &= x^3 \ln x - \int \frac{x^3}{x} \, dx \\ &= x^3 \ln x - \frac{1}{3}x^3 + C.\end{aligned}$$

**Example 6.1.4 An Integral from Differential Equations.** An important concept in the field of differential equations is that of the *Laplace transform*. In the process of computing the Laplace transform of  $\cos t$ , an integral similar to  $\int e^t \cos t \, dt$  appears. Determine this integral.

**Solution.** Here, it doesn't make much difference what we select for  $u$  or  $dv$ , since both  $e^t$  and  $\cos t$  are easy to differentiate and integrate, while neither becomes simpler. But since we're integrating a product, we'll try to use integration by parts anyway and hope for the best. To that end, let  $u = e^t$  and  $dv = \cos t$ . Then  $du = e^t \, dt$  and  $v = \sin t$ , which gives

$$\begin{aligned}\int e^t \cos t \, dt &= uv - \int v \, du \\ &= e^t \sin t - \int e^t \sin t \, dt.\end{aligned}$$

As expected, the new integral isn't any simpler than the one that we started with. However, it's not any more *complicated* either. So we'll try integration by parts on the new integral as well, with  $u = e^t$ ,  $du = e^t \, dt$ ,  $dv = \sin t \, dt$  and  $v = -\cos t$ . Then

$$\int e^t \sin t \, dt = -e^t \cos t + \int e^t \cos t \, dt.$$

So we have

$$\begin{aligned}\int e^t \cos t \, dt &= e^t \sin t - (-e^t \cos t + \int e^t \cos t \, dt) \\ &= e^t \sin t + e^t \cos t - \int e^t \cos t \, dt\end{aligned}$$

There is something very strange about this equation: *the same integral appears on both sides*. This is no longer a calculus problem, it's an algebra problem! Solving for the integral algebraically (and including the arbitrary constant!) gives

$$\int e^t \cos t \, dt = \frac{e^t \sin t + e^t \cos t}{2} + C.$$

Integrals with integrands of the form  $P(x)e^x$ ,  $P(x)\sin x$  or  $P(x)\cos x$  (where  $P(x)$  is a polynomial) can be successfully evaluated using the *tabular method*.

**Example 6.1.5 The Tabular Method.** Evaluate  $\int (3x^2 - 2x) \cos 3x \, dx$ .

Integration by parts also applies to definite integrals. You only need to remember to carry the limits through.

**Example 6.1.6 Sign of a Sine.** Let  $f(x) = x^2 \sin x$ . Does  $f(x)$  tend to be positive or negative along  $[0, 2\pi]$ ?

**Solution.** We need to check the sign of  $\int_0^{2\pi} x^2 \sin x \, dx$ . Using the tabular method, we get

$$\int_0^{2\pi} x^2 \sin x \, dx = [-x^2 \cos x + 2x \sin x + 2 \cos x]_0^{2\pi} = -4\pi^2 + 4\pi - 2.$$

Since this is negative,  $f(x)$  tends to be negative along  $[0, 2\pi]$ .

## 6.2 Trigonometric Integrals and Substitution

In this section we determine how to evaluate integrals involving powers of trigonometric functions (such as  $\sin$ ,  $\sec$  and others), as well as integrals involving certain radicals. Much of our work will follow from the important identity  $\sin^2 x + \cos^2 x = 1$ , as well as the *half-angle formulas*

$$\begin{aligned}\sin^2 x &= \frac{1 - \cos 2x}{2} \\ \cos^2 x &= \frac{1 + \cos 2x}{2}\end{aligned}$$

### 6.2.1 Integrals Involving Powers of Sine and Cosine

**Example 6.2.1 Integrating an Odd Power of Cosine.** Determine  $\int_0^{\frac{\pi}{6}} \cos^5(3x) \, dx$ .

**Solution.** Although it's not obvious, we can actually solve this integral with substitution. First, we'll separate a power of  $\cos 3x$  from the rest in order to act as our  $du$ :

$$\int_0^{\frac{\pi}{6}} \cos^4(3x) \cos 3x \, dx.$$

Now, if we want  $\cos 3x$  to act as  $du$  in a substitution, then we need  $u$  to look like  $\sin 3x$ . We can do this by using the Pythagorean Identity and writing  $\cos^4(3x) = (1 - \sin^2(3x))^2$ .

Now we'll set  $u = \sin 3x$ , which gives  $du = 3 \cos 3x \, dx$  and

$$\begin{aligned}\int_0^{\frac{\pi}{6}} \cos^4(3x) \cos 3x \, dx &= \int_0^{\frac{\pi}{6}} (1 - \sin^2(3x))^2 \cos 3x \, dx \\ &= \frac{1}{3} \int_0^{\frac{1}{2}} (1 - u^2)^2 \, du \\ &= \int_0^{\frac{1}{2}} (1 - 2u^2 + u^4) \, du,\end{aligned}$$

which we can finish integrating without much trouble.

The trick we used above only worked because we had an odd power of cosine. In general, when integrating  $\int \sin^m x \cos^n x \, dx$ , you want to separate a factor from an odd power (if there is one). Then finish by using substitution.

**Example 6.2.2 Odd Powers of Sine and Cosine.** Determine  $\int \sin^3 t \cos^5 t \, dt$ .

**Solution.** Remember: we want to separate a factor off of an odd power if there is one. Thankfully, we have two! We'll pick sine this time:

$$\int \sin^3 t \cos^5 t \, dt = \int \sin^2 t \cos^5 t \sin t \, dt$$



$$\begin{aligned}
&= \int (1 - \cos^2 t) \cos^5 t \sin t \, dt \\
&= - \int (1 - u^2) u^5 \, du \\
&= - \int (u^5 - u^7) \, du.
\end{aligned}$$

At this point, the integral isn't too difficult to complete.

If integrating even powers of sine and cosine, the algebra gets a little worse since the substitution trick applied above no longer works.

**Example 6.2.3 Integrating Even Powers of Cosine.** The integral  $\int_0^{\frac{\pi}{2}} \cos^2 \theta \, d\theta$  appears when finding the area of the circle using integration. Find the value of this integral.

**Solution.** Since we have an even power to work with, we'll try the half-angle formula. This lets us reduce the power on the cosine, at the expense of multiplying  $\theta$  by 2, which greatly works in our favor:

$$\begin{aligned}
\int_0^{\frac{\pi}{2}} \cos^2 \theta \, d\theta &= \int_0^1 \frac{1 + \cos 2\theta}{2} \, d\theta \\
&= \frac{1}{2} \int_0^{\frac{\pi}{2}} (1 + \cos 2\theta) \, d\theta \\
&= \frac{1}{2} \left[ \theta + \frac{1}{2} \sin 2\theta \right]_0^{\frac{\pi}{2}} \\
&= \frac{\pi}{4}.
\end{aligned}$$

## 6.2.2 Integrals Involving Powers of Secant and Tangent

In contrast to those involving sine and cosine, all of the integrals here involving secant and tangent are solvable using substitution. Generally, we'll attack these problems by separating a factor of  $\sec^2 x$  (in which case we use  $u = \tan x$ ) or a factor of  $\sec x \tan x$  (in which case we use  $u = \sec x$ ). This will often be done using the *Pythagorean identity*

$$\tan^2 x = \sec^2 x - 1$$

We will occasionally need to make use of two integral formulas:

$$\begin{aligned}
\int \sec x \, dx &= \ln |\sec x + \tan x| + C \\
\int \tan x \, dx &= \ln |\sec x| + C
\end{aligned}$$

**Example 6.2.4 Integrating a Power of Tangent.** Determine  $\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \tan^6 w \, dw$ .

**Solution.** Our goal is to separate a factor of either  $\sec^2 w$  or  $\sec w \tan w$  to serve as our  $du$ . Since neither appears, we'll use the Pythagorean identity to introduce a  $\sec^2 w$  into our integral:

$$\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \tan^6 w \, dw = \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \tan^4 w \sec^2 w \, dw - \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \tan^4 w \, dw.$$

At this point the first integral can be solved using substitution, but the second integral needs to be rewritten again. However, the same trick we applied

before works here too: just use the Pythagorean identity to introduce a factor of  $\sec^2 w$ . This gives

$$\begin{aligned}\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \tan^6 w \, dw &= \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \tan^4 w \sec^2 w \, dw - \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \tan^2 w \sec^2 w \, dw + \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \tan^2 w \, dw \\ &= \int_{-1}^1 u^4 \, du - \int_{-1}^1 u^2 \, du + \int_{-1}^1 du - \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} dw \\ &= \frac{2}{5} - \frac{2}{3} + 2 - \frac{\pi}{2}.\end{aligned}$$

So

$$\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \tan^6 w \, dw = \frac{2}{5} - \frac{2}{3} + 2 - \frac{\pi}{2}.$$

**Example 6.2.5 Integrating a Product of Tangent and Secant.** Determine  $\int \tan x \sec^3 x \, dx$

**Solution.** Here, we see that we can separate a factor of  $\tan x \sec x$  from the integrand, so we'll do that and use  $u = \sec x$ :

$$\begin{aligned}\int \tan x \sec^3 x \, dx &= \int \sec^2 x \sec x \tan x \, dx \\ &= \int u^2 \, du \\ &= \frac{\sec^3 x}{3} + C.\end{aligned}$$

In general, try to follow the guidelines below when integrating products of secant and tangent:

- If the power of secant is *even*, then separate a factor of  $\sec^2 x$  and use  $u = \tan x$ , along with the Pythagorean identity if necessary.
- If the power of tangent is *odd*, then separate a factor of  $\sec x \tan x$  and use  $u = \sec x$ , along with the Pythagorean identity if necessary.

Now, for an example that does not fall into the above guidelines.

**Example 6.2.6 Integrating Secant Cubed.** Determine  $\int \sec^3 x \, dx$ .

**Solution.** We're dealing with an odd power of secant, so splitting off a factor of  $\sec^2 x$  won't help here. We also lack a  $\tan x$  term, so splitting off a factor of  $\sec x \tan x$  is tricky as well. We'll try integrating by parts instead, with

$$\begin{aligned}u &= \sec x & du &= \sec x \tan x \, dx \\ dv &= \sec^2 x \, dx & v &= \tan x\end{aligned}$$

This gives

$$\begin{aligned}\int \sec^3 x \, dx &= \sec x \tan x - \int \sec x \tan^2 x \, dx \\ &= \sec x \tan x - \int \sec^3 x \, dx + \int \sec x \, dx \\ &= \sec x \tan x - \int \sec^3 x \, dx + \ln |\sec x + \tan x|\end{aligned}$$

Therefore

$$\int \sec^3 x \, dx = \frac{1}{2}(\sec x \tan x + \ln |\sec x + \tan x|) + C.$$

### 6.2.3 Trigonometric Substitutions

An ellipse with horizontal axis length  $a$  and vertical axis length  $b$  centered at the origin is given by

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

If we want to find the area of this ellipse, then we can compute the following integral and multiply it by 4:

$$\int_0^a \sqrt{b^2 - \left(\frac{b}{a}\right)^2 x^2} dx = \int_0^a \frac{b}{a} \sqrt{a^2 - x^2} dx.$$

This integral is difficult to solve due to the square root in the integrand; it's tough to imagine what an antiderivative might be. So what we'll try to do is to use a "reverse substitution" that simplifies the square root.

To do so, notice that the Pythagorean identity gives  $\cos^2 \theta = 1 - \sin^2 \theta$ . Taking the square root of the left hand side is much easier here than taking a root of the right hand side. So we'll try that substitution here: let

$$\begin{aligned} x &= a \sin \theta \\ dx &= a \cos \theta d\theta \end{aligned}$$

Then

$$\begin{aligned} \int_0^a \sqrt{a^2 - x^2} dx &= \int_{x=0}^{x=a} \sqrt{a^2 - a^2 \sin^2 \theta} a \cos \theta d\theta \\ &= a^2 \int_{x=0}^{x=a} \cos^2 \theta d\theta \end{aligned}$$

Before we go further, let's figure out what our new limits are. Since  $x = a \sin \theta$ , we have  $\theta = \sin^{-1} \frac{x}{a}$ , and so  $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$ . Hence we need our limits for  $\theta$  to be  $\theta = 0$  and  $\theta = \frac{\pi}{2}$ , and so we get

$$\int_0^a \sqrt{a^2 - x^2} dx = a^2 \int_0^{\frac{\pi}{2}} \cos^2 \theta d\theta.$$

We've seen how to integrate a function like this in [Example 6.2.3](#). Hence the area of this ellipse is

$$4 \int_0^a \frac{b}{a} \sqrt{a^2 - x^2} dx = 4ab \int_0^{\frac{\pi}{2}} \cos^2 \theta d\theta = ab\pi.$$

This method is known as **trigonometric substitution**, and is useful for simplifying radicals. [Table 6.2.7](#) shows the various substitutions that can be used.

radical	substitution	identity
$\sqrt{a^2 - x^2}$	$x = a \sin \theta$	$\cos^2 \theta = 1 - \sin^2 \theta$
$\sqrt{x^2 - a^2}$	$x = a \sec \theta$	$\tan^2 \theta = \sec^2 \theta - 1$
$\sqrt{a^2 + x^2}$	$x = a \tan \theta$	$\sec^2 \theta = 1 + \tan^2 \theta$

**Table 6.2.7:** Trigonometric substitutions

**Example 6.2.8 An Exponential Integral.** Determine  $\int \frac{e^t}{(4 + e^{2t})^{3/2}} dt$ .

This is adapted from *Thomas' Calculus*, 11<sup>th</sup> edition, exercise 30 on page 591.

**Solution.** Even though this integral involves only exponentials, we can simplify it using trigonometric substitution due to the presence of the radical. First, we'll use  $u$ -substitution to rewrite the integral into a form more suitable for trigonometric substitution: with  $u = e^t$  and  $du = e^t dt$ , we get

$$\int \frac{e^t}{(4 + e^{2t})^{3/2}} dt = \int \frac{du}{(4 + u^2)^{3/2}}.$$

The appearance of  $4 + u^2$  in the denominator suggests that we use a tangent substitution, namely,  $u = 2 \tan \theta$ . Then  $du = 2 \sec^2 \theta d\theta$ , and we get

$$\begin{aligned} \int \frac{du}{(4 + u^2)^{3/2}} &= \int \frac{2 \sec^2 \theta}{4^{3/2}(1 + \tan^2 \theta)^{3/2}} d\theta \\ &= \frac{1}{4} \int \cos \theta d\theta \\ &= \frac{1}{4} \sin \theta + C \\ &= \frac{1}{4} \sin \arctan \frac{u}{2} + C \\ &= \frac{1}{4} \sin \arctan \frac{e^t}{2} + C \end{aligned}$$

Using triangles, we can simplify this further to get

$$\int \frac{e^t}{(4 + e^{2t})^{3/2}} dt = \frac{e^t}{4\sqrt{4 + e^{2t}}} + C.$$

## 6.3 Partial Fractions

At this point we have tools available to us that allow us to compute integrals of products (Section 6.1), and trigonometric functions and certain radicals (Section 6.2). Now we'll move on to a tool that will help us integrate rational functions (quotients of polynomials). The idea behind this method is to write a rational function, such as  $\frac{1}{x^2 - 1}$ , in terms of simpler fractions. We'll demonstrate by example.

**Example 6.3.1 Integrating a Rational Function.** Let

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

Find  $\int f(x) dx$ .

**Solution.** We could try secant substitution, but that we'll instead focus on simplifying this fraction. First, factor the denominator:

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

Now, assume that we can find constants  $A_1, A_2$  such that

$$\frac{1}{x^2 - 1} = \frac{A_1}{x - 1} + \frac{A_2}{x + 1}.$$

If we could find such constants, we could integrate the right hand side instead of the left hand side. There are a few ways to find these constants, but the easiest might be to just clear fractions by multiplying both sides by  $x^2 - 1$ :

$$1 = A_1(x + 1) + A_2(x - 1). \quad (6.3.1)$$

So  $A_1, A_2$  need to satisfy this equation *for all possible values of  $x$* . We'll pick values for  $x$  that make it easy to find  $A_1$  and  $A_2$ . If  $x = -1$ , then (6.3.1) becomes

$$1 = -2A_2 \Rightarrow A_2 = -\frac{1}{2}.$$

Similarly,  $A_1 = \frac{1}{2}$ . Therefore

$$\begin{aligned} \int \frac{1}{x^2 - 1} dx &= \int \left( \frac{1}{2} \frac{1}{x - 1} - \frac{1}{2} \frac{1}{x + 1} \right) dx \\ &= \frac{1}{2} \ln \left| \frac{x - 1}{x + 1} \right| + C. \end{aligned}$$

Finding partial fraction decompositions as in [Example 6.3.1](#) isn't too bad as long as the denominator can be factored into distinct (i.e. non-repeated) *linear factors*. The situation is more complicated if the factors are repeated or are quadratics (or both).

**Example 6.3.2 Finding a Partial Fraction Decomposition with Repeated Powers.** Find the partial fraction decomposition for  $\frac{1}{(x^2 - 1)^2}$ .

**Solution.** Since  $\frac{1}{(x^2 - 1)^2} = \frac{1}{(x - 1)^2(x + 1)^2}$ , it's tempting to write

$$\frac{1}{(x^2 - 1)^2} = \frac{A_1}{x - 1} + \frac{A_2}{x + 1}.$$

Unfortunately, doing so gives  $A_1 = A_2 = 0$ , which obviously won't work. If we instead use

$$\frac{1}{(x^2 - 1)^2} = \frac{A_1}{(x - 1)^2} + \frac{A_2}{(x + 1)^2},$$

we get  $A_1 = A_2 = \frac{1}{4}$ . However,

$$\frac{1}{(x^2 - 1)^2} \neq \frac{\frac{1}{4}}{(x - 1)^2} + \frac{\frac{1}{4}}{(x + 1)^2}.$$

What we need to do is assume

$$\frac{1}{(x^2 - 1)^2} = \frac{A_1}{x - 1} + \frac{A_2}{x + 1} + \frac{A_3}{(x - 1)^2} + \frac{A_4}{(x + 1)^2}.$$

If we don't account for *each* power, we're not guaranteed a partial fraction decomposition. Now let's clear fractions:

$$1 = A_1(x - 1)(x + 1)^2 + A_2(x + 1)(x - 1)^2 + A_3(x + 1)^2 + A_4(x - 1)^2. \quad (6.3.2)$$

Quickly, using  $x = -1$  we find that  $A_4 = \frac{1}{4}$ , and using  $x = 1$  we find that  $A_3 = \frac{1}{4}$ .

At this point there are no nice values left to plug into the right hand side of (6.3.2), so we need to find  $A_1$  and  $A_2$  another way. One possibility is to differentiate both sides to make the right hand side simpler, which gives

$$0 = A_1(x + 1)^2 + 2A_1(x^2 - 1) + A_2(x - 1)^2 + 2A_2(x^2 - 1) + 2A_3(x + 1) + 2A_4(x - 1). \quad (6.3.3)$$

Now, entering  $x = -1$  into (6.3.3) gives

$$0 = 4A_2 - 4A_4$$

and so  $A_2 = A_4 = \frac{1}{4}$ . Similarly,  $A_1 = -A_3 = -\frac{1}{4}$ . Therefore

$$\frac{1}{(x^2 - 1)^2} = \frac{1}{4} \left[ -\frac{1}{x-1} + \frac{1}{x+1} + \frac{1}{(x-1)^2} + \frac{1}{(x+1)^2} \right].$$

In order to find partial fraction decompositions when the denominator contains a quadratic, we must allow for the possibility of the numerator being *linear*.

**Example 6.3.3 Partial Fraction Decomposition with a Quadratic.** Determine

$$\int \frac{5x^2 + x^3}{x^5 + 3x^3} dx.$$

**Solution.** First, we can simplify:

$$\frac{x^3 + 5x^2}{x^5 + 3x^3} = \frac{x + 5}{x(x^2 + 3)}.$$

Now that we've factored the denominator, we can attempt to find the partial fraction decomposition. So assume that

$$\frac{x + 5}{x(x^2 + 3)} = \frac{A_1}{x} + \frac{A_2x + A_3}{x^2 + 3}$$

Note the linear term  $A_2x + A_3$  above the quadratic!

The rest of this computation goes like it did before. If we clear fractions, we get

$$x + 5 = A_1(x^2 + 3) + A_2x^2 + A_3x.$$

If we plug in  $x = 0$ , we obtain  $A_1 = \frac{5}{3}$ . At this point, we could plug in  $x = \sqrt{3}i$  (which would be okay!) or differentiate both sides, but it might be easier to take the following approach: expand the right hand side and collect like terms to get

$$x + 5 = (A_1 + A_2)x^2 + A_3x + 3A_1.$$

For the right hand side to equal the left hand side,  $A_1 + A_2$  *must be zero*. Hence  $A_2 = -A_1 = -\frac{5}{3}$ . Similarly,  $A_3$  *must equal one*. Therefore

$$\begin{aligned} \int \frac{x^3 + 5x^2}{x^5 + 3x^3} dx &= \int \left[ \frac{5/3}{x} + \frac{1 - \frac{5}{3}x}{x^2 + 3} \right] dx \\ &= \frac{5}{3} \int \frac{1}{x} dx + \int \frac{1}{x^2 + 3} dx - \frac{5}{3} \int \frac{x}{x^2 + 3} dx \\ &= \frac{5}{3} \ln|x| + \frac{1}{\sqrt{3}} \tan^{-1} \frac{x}{\sqrt{3}} - \frac{5}{6} \ln(x^2 + 3). \end{aligned}$$

Comparing like terms as in Example 6.3.3 is a useful technique even beyond partial fractions.

The method of partial fractions works *as long as the degree of the numerator is strictly less than that of the denominator*.

**Algorithm 6.3.4 Partial Fractions Algorithm.** To find the partial fraction decomposition for the rational function  $\frac{P(x)}{Q(x)}$ , do the following:

1. Simplify  $\frac{P(x)}{Q(x)}$  as much as possible, using long division if necessary when  $\deg P(x) \geq \deg Q(x)$ .
2. Factor  $Q(x)$ .
3. Based on the factors of  $Q(x)$ , determine the proper form of the partial fraction decomposition.
4. Clear fractions.
5. Find the unknown constants of the decomposition by plugging values for  $x$  that simplify the equation, differentiating both sides, and/or comparing like terms.

**Example 6.3.5 Using Long Division.** Determine  $\int \frac{2y^4}{y^3 - y^2 + y - 1} dy$ .

**Solution.** Before setting up partial fractions, we should simplify the integrand as much as possible. Here, we can use long division since the degree of the numerator is greater than or equal to the degree of the denominator. Doing so, we get

$$\frac{2y^4}{y^3 - y^2 + y - 1} = 2y + 2 + \frac{2}{y^3 - y^2 + y - 1}.$$

We can integrate the first two terms on the right hand side easily, but we'll use partial fractions to rewrite the fraction:

$$\begin{aligned} \frac{2}{y^3 - y^2 + y - 1} &= \frac{2}{y^2(y - 1) + (y - 1)} \\ &= \frac{2}{y^2(y - 1)} \\ &= \frac{A_1}{y} + \frac{A_2}{y^2} + \frac{A_3}{y - 1} \end{aligned}$$

Now clear fractions to get

$$2 = A_1y(y - 1) + A_2(y - 1) + A_3y^2.$$

Setting  $y = 0$  gives  $A_2 = -2$ , and setting  $y = 1$  gives  $A_3 = 2$ . To find  $A_1$ , we can expand the right hand side to write

$$2 = A_1y^2 - A_1y + A_2y - A_2 + A_3y^2.$$

Hence  $A_1 = -A_3 = -2$ .

Therefore

$$\begin{aligned} \int \frac{2y^4}{y^3 - y^2 + y - 1} dy &= \int \left[ 2y + 2 - \frac{2}{y} - \frac{2}{y^2} + \frac{2}{y - 1} \right] dy \\ &= y^2 + 2y - 2 \ln |y| + \frac{2}{y} + 2 \ln |y - 1| + C. \end{aligned}$$

The next example shows a substitution that may occasionally work when partial fractions fails.

**Example 6.3.6 Twisty Substitution.** Determine  $\int_{\frac{1}{3}}^3 \frac{\sqrt{x}}{x^2 + x} dx$ .

**Solution.** Once again we have a radical causing problems, and it's not in any of the three trigonometric substitution forms we can use either. We want it to go away, so we'll try just that: set  $u = \sqrt{x}$ , which gives  $du = \frac{1}{2\sqrt{x}} dx$ . Then

$$\int_{\frac{1}{3}}^3 \frac{\sqrt{x}}{x^2 + x} dx = \int_{\frac{1}{\sqrt{3}}}^{\sqrt{3}} \frac{2x}{x^2 + x} du.$$

Now, we use the fact that  $x = u^2$  to write

$$\int_{\frac{1}{\sqrt{3}}}^{\sqrt{3}} \frac{2x}{x^2 + x} du = \int_{\frac{1}{\sqrt{3}}}^{\sqrt{3}} \frac{2u^2}{u^4 + u^2} du$$

which simplifies to

$$\int_{\frac{1}{\sqrt{3}}}^{\sqrt{3}} \frac{2}{u^2 + 1} du = 2 \tan^{-1}(\sqrt{3}) - 2 \tan^{-1}\left(\frac{1}{\sqrt{3}}\right) = \frac{\pi}{3}.$$

## 6.4 Improper Integrals

(Note: this section corresponds to Section 6.6 of the text.) Consider the following problem: determine the smallest velocity required for an object to escape the earth's gravitational pull (i.e., the **escape velocity**). We can try to answer this problem by looking at the total work  $W$  needed to move this object. From one viewpoint, work is just the total change in kinetic energy. If the object is moving at escape velocity, say  $v_e$ , then all of its kinetic energy must have been converted into work:  $W = \frac{1}{2}m_O v_e^2$ , where  $m_O$  is the mass of the object.

As we'll see later, work is also the integral of force, so let's examine the force on the object. If we neglect air resistance and assume that the only force acting on the object is gravity, then the force exerted on the object is given by

$$F(r) = \frac{m_E m_O}{r^2} G,$$

where  $m_E$  is the mass of the Earth,  $m_O$  is still the mass of the object,  $r$  is the distance from the center of mass of the Earth to the center of mass of the object, and  $G$  is the gravitational constant. Now assume that the Earth is a sphere of radius  $R_0$ , and that the object has enough velocity to move it a distance of  $R_1$  units from the surface of the Earth. Then the total work done in moving the object this distance is

$$\int_{R_0}^{R_1} F(r) dr = m_E m_O G \int_{R_0}^{R_1} \frac{1}{r^2} dr.$$

Now let's think about should happen if the object is at escape velocity. In this case, theoretically at least, there should be no limit to the distance the object can travel if given enough time. So what's the total work done?

Let's try to compute  $m_E m_O G \int_{R_0}^{R_1} \frac{1}{r^2} dr$ . We get

$$\begin{aligned} m_E m_O G \int_{R_0}^{R_1} \frac{1}{r^2} dr &= -\frac{m_E m_O G}{r} \Big|_{R_0}^{R_1} \\ &= \frac{m_E m_O G}{R_0} - \frac{m_E m_O G}{R_1}. \end{aligned}$$

Note that this is still an underestimate for the total work theoretically done, since  $R_1$  can increase without bound. So let's send  $R_1$  to  $\infty$  and see what we get:

$$\lim_{R_1 \rightarrow \infty} \frac{m_E m_O G}{R_0} - \frac{m_E m_O G}{R_1} = \frac{m_E m_O G}{R_0}.$$

So, to summarize, the total work  $W$  should be

$$W = \lim_{R_1 \rightarrow \infty} \left( m_E m_O G \int_{R_0}^{R_1} \frac{1}{r^2} dr \right) = \frac{1}{2} m_O v_e^2$$



or just

$$W = \frac{m_E m_O G}{R_0} = \frac{1}{2} m_O v_e^2.$$

We can solve this equation for  $v_e$ , and so the escape velocity should be

$$v_e = \sqrt{\frac{2m_E G}{R_0}}.$$

As a neat bonus, this shows that the escape velocity only depends on  $G$ , the mass of the Earth  $m_E$  and the "radius" of the Earth  $R_0$ . Plugging in values for these figures, we get  $v_e \approx 11185$  meters per second, or 11.185 kilometers per second.

This is all very important obviously, but this is calculus, and what *we* really care about are the mathematical tools required to solve this problem. To do this, we basically had to figure out what the integral of  $\frac{1}{r^2}$  was over the interval  $[R_0, \infty)$ . This is our first example of an **improper integral**.

### 6.4.1 Improper Integrals over Infinite Intervals

First, a definition.

**Definition 6.4.1 Type 1 Improper Integral.** Let  $f(x)$  be some function, and let  $a$  be a constant. If  $\int_a^b f(x) dx$  exists for every  $b \geq a$ , then we define

$$\int_a^\infty f(x) dx = \lim_{b \rightarrow \infty} \int_a^b f(x) dx,$$

assuming this limit exists. Similarly, if  $\int_b^a f(x) dx$  exists for every  $b \leq a$ , then we define

$$\int_{-\infty}^a f(x) dx = \lim_{b \rightarrow -\infty} \int_b^a f(x) dx,$$

assuming this limit exists. These improper integrals are **convergent** if the corresponding limits exist and **divergent** otherwise.

Finally, if both  $\int_{-\infty}^a f(x) dx$  and  $\int_a^\infty f(x) dx$  are convergent, then we define

$$\int_{-\infty}^\infty f(x) dx = \int_{-\infty}^a f(x) dx + \int_a^\infty f(x) dx.$$

**Example 6.4.2 A Divergent Integral.** Determine if  $\int_1^\infty \frac{1}{x} dx$  is convergent or divergent.

**Solution.** By definition, we have

$$\begin{aligned} \int_1^\infty \frac{1}{x} dx &= \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x} dx \\ &= \lim_{b \rightarrow \infty} (\ln b - \ln 1) \\ &= \infty. \end{aligned}$$

Hence the integral is divergent.

In [Example 6.4.2](#), note that  $\frac{1}{x} \rightarrow 0$  as  $x \rightarrow \infty$ . However, it doesn't go to 0 *fast enough* for the integral to converge.

**Example 6.4.3.** Is  $\int_1^\infty \frac{1}{x^2} dx$  convergent?

The convergence of  $\int_1^\infty \frac{1}{x^p} dx$  depends on the value of  $p$ .

**Theorem 6.4.4 Integral  $p$ -test.** The integral  $\int_1^\infty \frac{1}{x^p} dx$  converges if and only if  $p > 1$ .

**Example 6.4.5 Area Under a Graph.** Find the area under  $f(x) = xe^{-x^2}$ .

**Solution.** The area under  $f(x)$  is given by  $\int_{-\infty}^\infty xe^{-x^2} dx$ . To evaluate this integral, we must split it into two improper integrals:

$$\begin{aligned}\int_{-\infty}^\infty xe^{-x^2} dx &= \int_{-\infty}^0 xe^{-x^2} dx + \int_0^\infty xe^{-x^2} dx \\ &= \lim_{a \rightarrow -\infty} \int_a^0 xe^{-x^2} dx + \lim_{b \rightarrow \infty} \int_0^b xe^{-x^2} dx\end{aligned}$$

At this point, we can use  $u$ -substitution to find each integral, using  $u = -x^2$  and  $du = -2x dx$ . So we get

$$\begin{aligned}\lim_{a \rightarrow -\infty} \int_a^0 xe^{-x^2} dx + \lim_{b \rightarrow \infty} \int_0^b xe^{-x^2} dx &= \lim_{a \rightarrow -\infty} \int_{-a^2}^0 \frac{-e^u}{2} du + \lim_{b \rightarrow \infty} \int_0^{-b^2} \frac{-e^u}{2} du \\ &= \lim_{a \rightarrow -\infty} \left[ -\frac{1}{2} + \frac{e^{-a^2}}{2} \right] + \lim_{b \rightarrow \infty} \left[ -\frac{e^{-b^2}}{2} + \frac{1}{2} \right] \\ &= \lim_{a \rightarrow -\infty} \frac{e^{-a^2}}{2} - \lim_{b \rightarrow \infty} \frac{e^{-b^2}}{2} \\ &= 0.\end{aligned}$$

Hence the area under  $f(x)$  is 0 (which might not be so surprising if you graph  $f(x)$ ).

**Example 6.4.6 Radioactive Decay.** A radioactive substance undergoes decay, and has mass  $m(t)$  at time  $t$  given by

$$m(t) = m_0 e^{-kt}$$

for some  $k > 0$ . The *expected lifetime* of a particle is given by

$$k \int_0^\infty te^{-kt} dt.$$

Find this value.

**Solution.** We can solve this using integration by parts. Doing so, we get

$$\begin{aligned}k \int_0^\infty te^{-kt} dt &= -\lim_{b \rightarrow \infty} \left[ te^{-kt} + \frac{1}{k} e^{-kt} \right]_0^b \\ &= \frac{1}{k}\end{aligned}$$

This can be verified using the code cell below.

```
from sympy import *
init_printing()

# We need to assume that k is positive for the integral to
# converge.
# This will also make SymPy assume that t is positive, but
# that's not a problem for us
# since we're integrating over [0,oo).
k,t = symbols('k_t', positive=True)
m = k*t*exp(-k*t)

m.integrate((t, 0, oo))
```

### 6.4.2 Improper Integrals with Discontinuous Integrands

The second type of improper integral we'll consider involves integrands with "divide-by-zero" problems.

**Definition 6.4.7 Type 2 Improper Integral.** If  $f(x)$  is continuous on  $[a, b)$  and is discontinuous at  $b$ , then

$$\int_a^b f(x) dx = \lim_{c \rightarrow b^-} \int_a^c f(x) dx$$

assuming this limit exists.

If  $f(x)$  is continuous on  $(a, b]$  and is discontinuous at  $a$ , then

$$\int_a^b f(x) dx = \lim_{c \rightarrow a^+} \int_c^b f(x) dx$$

assuming this limit exists.

If  $f(x)$  is at  $c$  where  $a < c < b$ , then

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$$

assuming these integrals are convergent.

**Example 6.4.8 Logarithmic Discontinuity.** Determine if  $\int_0^1 \ln x dx$  is convergent.

**Solution.** This integral is improper since  $\ln x$  is discontinuous at  $x = 0$ . To evaluate it, we'll use limits:

$$\begin{aligned} \int_0^1 \ln x dx &= \lim_{c \rightarrow 0^+} \int_c^1 \ln x dx \\ &= \lim_{c \rightarrow 0^+} [x \ln x - x]_c^1 \\ &= -1 - \lim_{c \rightarrow 0^+} c \ln c \\ &= -1, \end{aligned}$$

where we used the fact that  $c \ln c \rightarrow 0$  as  $c \rightarrow 0$  by L'Hopital's Rule.

**Example 6.4.9 Another Integral  $p$ -Test.** For what values of  $p$  is the integral  $\int_0^1 \frac{1}{\sqrt[p]{x}} dx$  convergent?

**Solution.** We could try to compute this as is, but we can save ourselves some work by making use of [Theorem 6.4.4](#). First, we'll use the substitution  $u = x^{-\frac{1}{p}}$ . Then  $du = -\frac{1}{p}x^{-\frac{1}{p}-1}$  and

$$\begin{aligned} u &\rightarrow \infty \quad \text{as } x \rightarrow 0^+ \\ u &= 1 \quad \text{if } x = 1. \end{aligned}$$

Therefore

$$\begin{aligned} \int_0^1 \frac{1}{\sqrt[p]{x}} dx &= \int_\infty^1 u [-pu^{-1-p}] du \\ &= p \int_1^\infty \frac{1}{u^p} du \end{aligned}$$

which we know converges if and only if  $p > 1$ . Hence the original integral converges if and only if  $p > 1$ .

### 6.4.3 Comparison Tests for Improper Integrals

For integrals, we have two comparison tests that allow us to determine if an improper integral converges by comparing it with a simpler integral. First, we'll look at the **direct comparison test**.

**Theorem 6.4.10 Direct Comparison Test.** *Suppose that  $f(x)$  and  $g(x)$  are continuous functions and that*

$$f(x) \geq g(x) \geq 0$$

*for  $x \geq a$ , for some real number  $a$ . Then if  $\int_a^\infty f(x) dx$  converges, so goes  $\int_a^\infty g(x) dx$ . Likewise, if  $\int_a^\infty g(x) dx$  diverges, then so does  $\int_a^\infty f(x) dx$ .*

**Example 6.4.11.** Does  $\int_1^\infty \frac{1}{\sqrt{e^x+x}} dx$  converge?

**Solution.** This integrand looks pretty awful. However, we can say that

$$\frac{1}{\sqrt{e^x+x}} \leq \frac{1}{\sqrt{e^x}} = e^{-\frac{x}{2}}.$$

Since

$$\int_1^\infty e^{-\frac{x}{2}} dx = 2\sqrt{e},$$

[Theorem 6.4.10](#) implies that the original integral must converge as well. Note that we don't know *what* it converges to, only that it does.

**Example 6.4.12 The Gaussian Integral.** Show that the integral  $\int_0^\infty e^{-x^2} dx$  is convergent.

**Solution.** We can do this by breaking the integral into two parts:  $\int_0^\infty e^{-x^2} dx = \int_0^1 e^{-x^2} dx + \int_1^\infty e^{-x^2} dx$ . If  $x$  is between 0 and 1, then  $e^{-x^2} \leq e^0 = 1$ . On the other hand, if  $1 \leq x < \infty$ , then  $e^{-x^2} \leq e^{-x}$ . Therefore

$$\begin{aligned} \int_0^\infty e^{-x^2} dx &= \int_0^\infty e^{-x^2} dx = \int_0^1 e^{-x^2} dx + \int_1^\infty e^{-x^2} dx \\ &\leq \int_0^1 dx + \int_1^\infty e^{-x} dx \\ &= 1 + \frac{1}{e} \\ &< \infty. \end{aligned}$$

Hence the original integral must also be convergent. It's value is actually known, and SymPy can be used to compute the integral, as below.

```
# Computing the Gaussian with SymPy.
import sympy as sp
sp.init_printing()

x = sp.Symbol('x')
sp.integrate(sp.exp(-x**2), (x, 0, oo))
```

To use [Theorem 6.4.10](#) successfully, we need to choose a simpler function that's related to the integrand and then use algebra to justify a specific inequality. It is sometimes more straightforward to apply the **limit comparison test**.

**Theorem 6.4.13 Limit Comparison Test.** *Let  $f(x)$  and  $g(x)$  be positive, continuous functions on  $[a, \infty)$  for some constant  $a$ . If*

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)}$$

*exists and is positive, then the integrals*

$$\int_a^\infty f(x) \, dx \quad \text{and} \quad \int_a^\infty g(x) \, dx$$

*must both converge or both diverge.*

**Example 6.4.14.** Determine if  $\int_1^\infty \frac{3x^3 + 2x^5}{5x^7 + 4x^2} \, dx$  converges.

**Solution.** The integrand is awful again, but if  $x$  is very large then we can say that

$$\frac{3x^3 + 2x^5}{5x^7 + 4x^2} \approx \frac{1}{x^2}.$$

Since  $\int_1^\infty \frac{1}{x^2} \, dx$  converges, this suggests that maybe the original integral does as well. We can prove this using [Theorem 6.4.13](#).

In fact, we have

$$\lim_{x \rightarrow \infty} \frac{\frac{3x^3 + 2x^5}{5x^7 + 4x^2}}{\frac{1}{x^2}} = \frac{2}{5}.$$

Since this limit exists and is positive, then the original integral must also converge by the Limit Comparison Test.

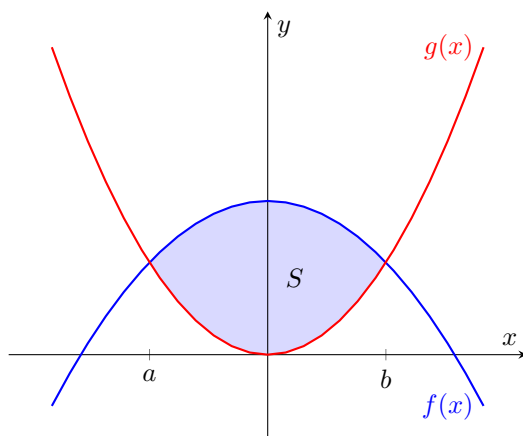


## Chapter 7

# Applications of Integration

### 7.1 Areas Between Curves

We already know that  $\int_a^b f(x) dx$  geometrically represents the area under  $f(x)$  from  $x = a$  to  $x = b$ . Now we'd like to try to extend the integral to other geometrical concepts such as surface areas and volumes. An easy generalization we can make right now is that of *areas between curves*, as in [Figure 7.1.1](#).



**Figure 7.1.1:** Area between curves. Adapted from [here](#).

It may not be too much of a stretch to say that the area  $S$  should be given by  $\int_a^b [f(x) - g(x)] dx$ , and indeed it is, but we will prove this using Riemann sums for the reason that Riemann sums appear *a lot* in applications. Furthermore, Riemann sums provide us with the formula we need to evaluate integrals numerically using libraries such as NumPy. So let's start by partitioning the interval  $[a, b]$  into  $n$  *subintervals*

$$[x_{i-1}, x_i] \quad \text{for } 0 \leq i \leq n,$$

where  $x_0 = a, x_n = b$  and

$$x_i - x_{i-1} = \Delta x = \frac{b - a}{n}.$$

Now we'll use the subintervals to break the region  $S$  up into rectangles to approximate the area. So choose a point  $x_i^*$  from each subinterval  $[x_{i-1}, x_i]$  and use it to determine the height of the corresponding rectangle:  $f(x_i^*) -$

$g(x_i^*)$ . Then the area is  $f(x_i^*) - g(x_i^*)\Delta x$ , and adding up the areas of all of the rectangles lets us approximate the area of  $S$ :

$$\text{area of } S \approx \sum_{i=1}^n [f(x_i^*) - g(x_i^*)]\Delta x \quad (7.1.1)$$

Now send  $n \rightarrow \infty$  (equivalently, send  $\Delta x \rightarrow 0$ ) to make the approximation exact:  $\text{area of } S = \lim_{\Delta x \rightarrow 0} \sum_{i=1}^n f(x_i^*) - g(x_i^*)\Delta x = \int_a^b [f(x) - g(x)] dx$ .

**Example 7.1.2 Area Between Parabolas.** Find the area between  $f(x) = x^2 + 1$  and  $g(x) = 2x^2$ .

**Solution.** If we graph  $f$  and  $g$ , we see that the area will be given by  $\int_a^b [(x^2 + 1) - 2x^2] dx$ , where  $a$  and  $b$  are the values of  $x$  where  $f(x) = g(x)$ . Using a bit of algebra, we get  $a = -1$  and  $b = 1$ . Therefore the area between the parabolas is

$$\begin{aligned} \int_{-1}^1 (1 - x^2) dx &= 2 \int_0^1 (1 - x^2) dx \\ &= \frac{4}{3}. \end{aligned}$$

Although we found an exact solution in [Example 7.1.2](#), we can also use this to test our NumPy skills. (7.1.1) shows us how we can use NumPy to find this quantity.

```
import numpy as np

# Endpoints of interval and number of subintervals
a = -1
b = 1
n = 10

# Quick way to define f(x) and g(x)
f = lambda x: x**2 + 1
g = lambda x: 2 * x**2

# Partition interval
X = np.linspace(a, b, n)

# Area of each rectangle
deltaX = (b - a) / n
area = (f(X) - g(X)) * deltaX

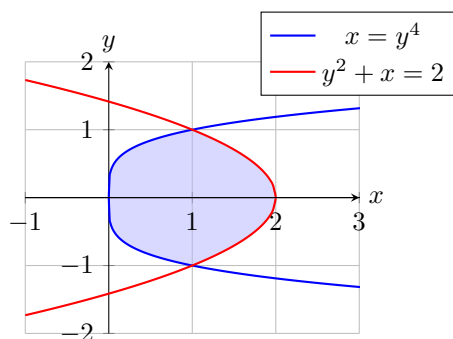
# Add areas together to get approximation
print(np.sum(area))
```

```
1.1851851851851853
```

**Example 7.1.3 Area with Changing Limits.** Find the area of the region bounded by  $x = y^4$  and  $y^2 + x = 2$ .

**Solution.** If we solve both equations for  $y$ , we get  $y = \pm\sqrt[4]{x}$  and  $y = \pm\sqrt{2-x}$ . We can then graph the region below:





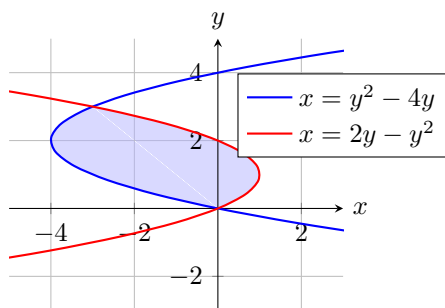
**Figure 7.1.4:** The region bounded by  $x = y^4$  and  $y^2 + x = 2$ .

To find the area between these curves, we'll need to compute *two* integrals: one for  $0 \leq x \leq 1$  and another for  $1 \leq x \leq 2$ . In particular, the area will be given by  $\int_0^1 [x^{\frac{1}{4}} - (-x^{\frac{1}{4}})] dx + \int_1^2 [\sqrt{2-x} - (-\sqrt{2-x})] dx$ . Doing so, we get  $\frac{44}{15}$  for the area between the curves.

Part of the reason why finding the area in [Example 7.1.3](#) was more complicated was because the curves did not represent functions of  $x$ . However, we can view each curve as representing a function of  $y$ , namely,  $x(y) = y^4$  and  $x(y) = 2 - y^2$ . Then the area can be found as  $\int_{-1}^1 [(2 - y^2) - y^4] dy = \frac{44}{15}$ . So it can be much easier if we find the area by integrating with respect to  $y$  as opposed to  $x$ .

**Example 7.1.5.** Find the area contained by the curves  $x = y^2 - 4y$  and  $x = 2y - y^2$ .

**Solution.** If we graph the region, we get Figure ??.



Finding the area by using integrals with respect to  $x$  would require setting up and solving three different integrals! So, not terrific. However, we can integrate with respect to  $y$  fairly easily. By finding where the curves intersect, we see that  $0 \leq y \leq 3$ . Hence the area should be

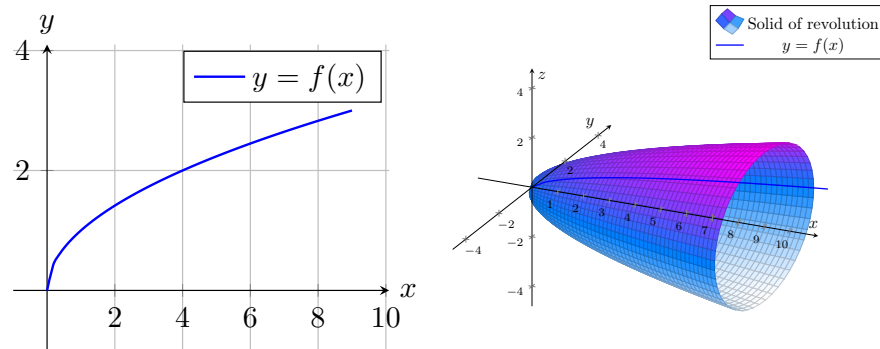
$$\begin{aligned} \int_0^3 [(2y - y^2) - (y^2 - 4y)] dy &= \int_0^3 [6y - 2y^2] dy \\ &= \left[ 3y^2 - \frac{2}{3}y^3 \right]_0^3 \\ &= 27 - 18 \end{aligned}$$

and so the area is just 9 units.

## 7.2 Volumes

Just as we can with areas, we can attempt to find volumes using integrals as well. At the moment we can't find volumes of general regions without first describing integrals in higher dimensions, which is the content of [Chapter 11](#). Until then, we'll need to restrict ourselves to so-called **solids of revolution**.

As an example, consider [Figure 7.2.1](#). To generate the solid of revolution on the right, we take the graph of  $y = f(x)$  and rotate it about the  $x$ -axis, creating a three-dimensional region.



**Figure 7.2.1:** Generating a solid of revolution.

To find the volume of such a region, the idea is to find the *area* of a general cross-section of the region, which by design should be a function of  $x$  alone, say  $A(x)$ . Then, roughly,  $A(x)\Delta x$  should represent the volume of a small cross-section of the solid, and so the integral  $\int_a^b A(x) dx$  should give us the volume of the region. This entire argument can be made precise using Riemann sums as in the previous section, and so we get the following: the solid  $S$  from  $x = a$  to  $x = b$  whose cross-sections (with respect to the  $x$ -axis) have area  $A(x)$  has volume given by

$$V = \int_a^b A(x) dx. \quad (7.2.1)$$

**Example 7.2.2 Finding the Volume of a Sphere.** Find the volume of a sphere of radius  $r$ .

**Solution.** To begin, let's assume that the sphere is centered at the origin. Then each cross section perpendicular to the  $x$ -axis is just a circle with radius  $\sqrt{r^2 - x^2}$ . Hence each cross-sectional area is given by  $A(x) = \pi(r^2 - x^2)$ , which means the volume of the sphere should be

$$\begin{aligned} \int_{-r}^r \pi(r^2 - x^2) dx &= 2 \int_0^r \pi(r^2 - x^2) dx \\ &= 2 \left[ \pi r^2 x - \frac{\pi x^3}{3} \right]_0^r \\ &= 2 \left[ \pi r^3 - \frac{\pi r^3}{3} \right] \\ &= \frac{4}{3} \pi r^3. \end{aligned}$$

Again, the idea for finding the volumes of these regions is to determine the cross-sectional areas perpendicular to the  $x$ -axis (or whatever the axis of rotation happens to be) and then integrate.

**Example 7.2.3.** Find the volume of the region obtained by rotating the graph of  $y = \sqrt{x}$  from  $x = 1$  to  $x = 4$  about the line  $y = 2$ .

**Solution.** As is almost always the case, a good way to start is by graphing the region to get a rough idea of what it looks like. Doing so, we see that each of the cross-sections are circles. But this is to be expected, since the region is a solid of revolution. Given  $x$  between 1 and 4, the area of the corresponding cross-section is given by

$$A(x) = \pi(2 - \sqrt{x})^2.$$

Hence the volume of this region should be

$$\begin{aligned} \int_1^4 \pi(2 - \sqrt{x})^2 dx &= \pi \int_1^4 (2 - 4\sqrt{x} + x) dx \\ &= \pi \left[ 2x - \frac{8}{3}x^{\frac{3}{2}} + \frac{x^2}{2} \right]_1^4 \\ &= \pi \left[ \left( 8 - \frac{64}{3} + 8 \right) - \left( 2 - \frac{8}{3} + \frac{1}{2} \right) \right] \end{aligned}$$

The previous problems involved finding volumes by using disks since disks were natural cross-sections. The next example instead uses washers as cross-sections instead. Although the formula we get for the volume is slightly different, it's still essentially integrating the areas of the cross-sections.

**Example 7.2.4.** Find the volume of the solid obtained by rotating the region bounded by  $y = x$ ,  $y = \sqrt{x}$  and  $y = 1$  about the line  $x = 2$ .

**Solution.** Since the axis of rotation is  $x = 2$ , we want to look at cross-sections perpendicular to this line, i.e., parallel to the  $x$ -axis. Each such cross-section will look like a washer, and hence its area will involve the inner radius and outer radius of the washer. Since the radii are measured in terms of horizontal distance, we'll need to find  $x$  in terms of  $y$ . Doing so, we see that the inner radius is given by  $2 - y$  while the outer radius is given by  $2 - y^2$ . Therefore the volume of the region is given by

$$\begin{aligned} \int_0^1 [\pi(2 - y^2)^2 - \pi(2 - y)^2] dy &= \pi \int_0^1 [(4 - 4y^2 + y^4) - (4 - 2y + y^2)] dy \\ &= \pi \int_0^1 [y^4 - 5y^2 + 2y] dy \end{aligned}$$

## 7.3 Volumes by Cylindrical Shells

The method we used to find volumes in [Section 7.2](#) applies to solids that have cross-sections depending on  $x$  or  $y$  alone. Once we had a formula for this area, we could then integrate it to get the volume. However, not all shapes have such simple cross-sections, and so this method doesn't apply in general. In this section, we'll look at using cylindrical shells instead.

Therefore, the following fact will prove useful: the volume of a cylindrical shell of height  $h$ , inner radius  $r_1$  and outer radius  $r_2$  is given by

$$V = 2\pi \frac{r_1 + r_2}{2} h (r_2 - r_1) = 2\pi r h \Delta r \quad (7.3.1)$$

where  $r$  is the "average radius" and  $\Delta r$  is the thickness of the shell. If we can break a solid up into these cylindrical "cross-sections," then integrating these areas should again give us the volume of the solid. For shells that are perpendicular to the  $x$ -axis, we can typically replace  $r$  with  $x$  and  $h$  with  $f(x)$ , leading to an integral of the form

$$\int_a^b 2\pi x f(x) dx.$$

Finding areas using disks and washers usually involves finding disks/washers that are perpendicular to the axis of rotation. When finding volumes using shells, we'll typically use shells that are *parallel* to the axis of rotation.

One final note: this method, as in [Section 7.2](#), applies only to solids with a certain degree of symmetry. If you continue on to Calculus 3, the techniques you'll learn in [Chapter 11](#) will replace this method as they do the disk and washer methods from [Section 7.2](#). In particular, this method is a special case of the transformation to *cylindrical coordinates* presented in [Section 11.6](#).

**Example 7.3.1 Finding Volumes Using Cylindrical Shells.** Find the volume of the solid obtained by rotating the region bounded by  $y = x^2$ ,  $y = 0$  and  $x = 1$  about the  $y$ -axis.

**Solution.** The axis of rotation is the  $y$ -axis, so each shell should be perpendicular to this axis as well. Drawing a typical shell, we see that the radius at  $x$  should be  $x$  while the height is  $f(x) = x^2$ . So the volume of this region should be

$$\int_0^1 2\pi x(x^2) dx = \frac{\pi}{2}.$$

It should be noted that the volume in [Example 7.3.1](#) could also have been found using washers instead of shells. If you can, you should use disks or washers before using shells, since integrals derived from using shells are typically more complicated.

**Example 7.3.2.** Find the volume of the solid obtained by rotating the region bounded by  $x = 0$ ,  $y = 4x - x^2$  and  $y = 8x - 2x^2$  about the line  $x = -2$ .

**Solution.** A good first step, as always, is to plot the region. Since  $y = 4x - x^2$  and  $y = 8x - 2x^2$  give downward opening parabolas, we can quickly sketch a rough graph by finding the  $x$ -coordinates of the vertices, which occur at  $x = 2$ . Finding the corresponding  $y$  values, and noticing that each parabola intersects the  $x$ -axis at  $x = 0, 4$ , gives a pretty good picture of the graph.

If we tried to find the volume using disks/washers, we'd need to solve  $y = 4x - x^2$  and  $y = 8x - 2x^2$  for  $x$  since each cross-section would be parallel to the  $x$ -axis. This is doable using the quadratic formula, but the algebra would get very bad very quickly once we start integrating. So we'll try shells instead.

A typical shell for this region has radius  $x - (-2) = x + 2$  and height  $(8x - 2x^2) - (4x - x^2) = 4x - x^2$ . Hence the volume is given by

$$\begin{aligned} \int_0^4 2\pi(x+2)(4x-x^2) dx &= 4\pi \int_0^2 (x+2)(2x-x^2) dx \\ &= 4\pi \int_0^2 [4x - x^3] dx \end{aligned}$$

$$\begin{aligned} &= 4\pi \left[ 2x^2 - \frac{x^4}{4} \right]_0^2 \\ &= 16\pi \end{aligned}$$

## MyOpenMath problems

These problems are only intended for practice. They are *not* required.

**Exercises for Section 7.1** These exercises deal with finding areas between curves.

1.

2.

3.

4.

5.

**Exercises for Section 7.2**

6.

7.

8.

9.



## Chapter 8

# Parametric Equations and Polar Coordinates

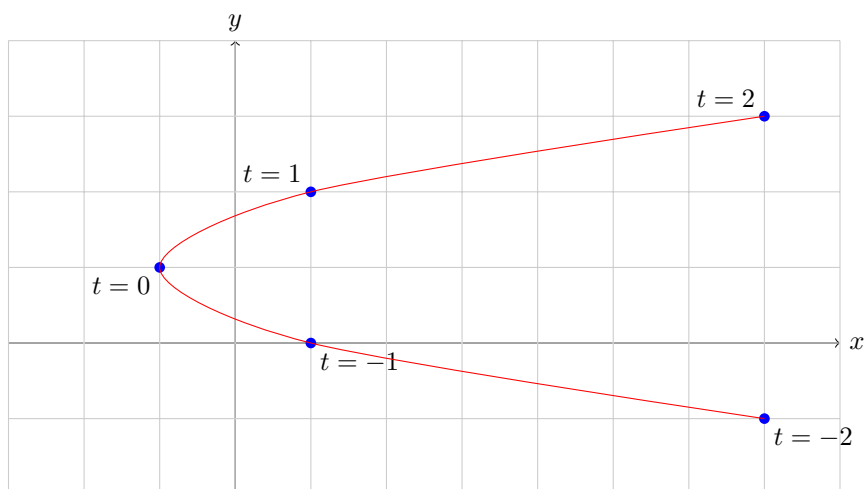
Up to this chapter we've represented curves in the  $xy$ -plane using Cartesian (i.e.  $xy$ ) coordinates. What we'll do now is introduce a new method of describing curves in the plane by way of parametric equations.

### 8.1 Parametric Equations

We can use **parametric equations** to describe **parametric curves**. For example, the parametric equations

$$\begin{aligned}x &= 2t^2 - 1 \\y &= t + 1 \\-2 &\leq t \leq 2\end{aligned}$$

produce the following parametric curve in the  $xy$ -plane.



**Figure 8.1.1:** The graph of  $x = 2t^2 - 1$ ,  $y = t + 1$  for  $-2 \leq t \leq 2$ .

Plotting points and plugging them into parametric equations can be tedious. Another way to describe a parametric curve is to eliminate parameters.

**Example 8.1.2 Eliminating parameters.** Consider the parametric equations

$$\begin{aligned}x &= 2t^2 - 1 \\ y &= t + 1\end{aligned}$$

Then we can solve for  $t$  in the second equation to get  $t = y - 1$ . If we plug this into the first equation, we get an equation in  $x$  and  $y$  alone:

$$x = 2(y - 1)^2 - 1.$$

We can see from this equation that the original parametric equations should trace out a rightward opening parabola. Note however that this equation tells us nothing about which portion of the parabola is traced out.

SUGGESTED PROBLEMS: 1, 7, 11

## 8.2 Calculus and Parametric Curves

### 8.2.1 Tangent Lines

We have an idea of how to represent curves in the  $xy$ -plane using parametric equations. What we'd like to do now is to figure out how to do calculus with parametric equations. To start, we'll find the slope of a parametric curve at some value  $t$ . Suppose that

$$\begin{aligned}x &= f(t) \\ y &= g(t).\end{aligned}$$

Then we can find the slope  $\frac{dy}{dx}$  using the chain rule:

$$\frac{dy}{dt} = \frac{dy}{dx} \frac{dx}{dt}$$

and so  $\frac{dy}{dx} = \frac{dy}{dt} / \frac{dx}{dt} = g'(t)/f'(t)$ , assuming that  $g'(t) \neq 0$ .

### 8.2.2 Area Under Parametric Curves

We can also find areas by using  $u$ -substitution and the Fundamental Theorem of Calculus. If  $y = F(x)$ , then the area under the curve  $y = F(x)$  from  $x = a$  to  $x = b$  is given by  $\int_a^b F(x) dx$ . If we instead have parametric equations

$$\begin{aligned}x &= f(t) \\ y &= g(t),\end{aligned}$$

then  $y = F(x) = F(f(t)) = g(t)$  and  $dx = f'(t) dt$ . If the curve goes from  $x = a$  to  $x = b$  exactly once as  $t$  goes from  $\alpha$  to  $\beta$ , then the area under the curve is given by

$$\int_{\alpha}^{\beta} g(t) f'(t) dt.$$

**Example 8.2.1 Area of a Circle.** Suppose we want to find the area of a circle of radius  $r$ . We can do so by parameterizing such a circle with

$$x = r \cos t$$



$$y = r \sin t$$

$$0 \leq t \leq 2\pi.$$

Then the area of the circle is given by

$$\begin{aligned} 2 \int_{\pi}^0 -r^2 \sin^2 t \, dt &= -2r^2 \int_{\pi}^0 \frac{1 - \cos 2t}{2} \, dt \\ &= -r^2 \left[ t - \frac{1}{2} \sin 2t \right]_{\pi}^0 \\ &= \pi r^2. \end{aligned}$$

One thing to watch out for in the previous calculation is that the limits went from  $\pi$  to 0. This was because of the particular parameterization we chose.

### 8.2.3 Arc Length

Now we move on to finding lengths of parametric curves. We won't derive the formula here, but it essentially follows from the Pythagorean theorem. We'll see a motivation of the formula later using vector calculus. But for now, suppose we have

$$x = f(t)$$

$$y = g(t).$$

Then the length of the associated parametric curve from  $t = \alpha$  to  $t = \beta$  is given by

$$\int_{\alpha}^{\beta} \sqrt{f'(t)^2 + g'(t)^2} \, dt = \int_{\alpha}^{\beta} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} \, dt.$$

**Example 8.2.2 Arc Length of a Parametric Curve.** Let  $x = e^t + e^{-t}$  and  $y = 5 - 2t$  for  $0 \leq t \leq 3$ . Then the length of this curve is given by

$$\begin{aligned} \int_0^3 \sqrt{(e^t - e^{-t})^2 + 4} \, dt &= \int_0^3 \sqrt{e^{2t} - 2 + e^{-2t} + 4} \, dt \\ &= \int_0^3 \sqrt{e^{2t} + 2 + e^{-2t}} \, dt \\ &= \int_0^3 \sqrt{(e^t + e^{-t})^2} \, dt \\ &= \int_0^3 (e^t + e^{-t}) \, dt \\ &= e^3 - e^{-3}. \end{aligned}$$

SUGGESTED PROBLEMS: 1, 3, 13, 37

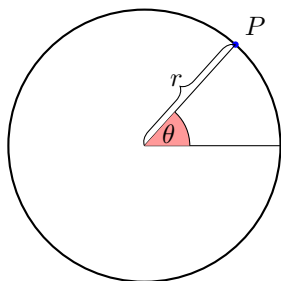
## 8.3 Polar Coordinates

### 8.3.1 Introducing Polar Coordinates

We typically use **Cartesian** or **rectangular** coordinates to plot points. However, this can lead to issues if we have a graph that isn't rectangular. For example, the circle has a simple geometric description based on its center and

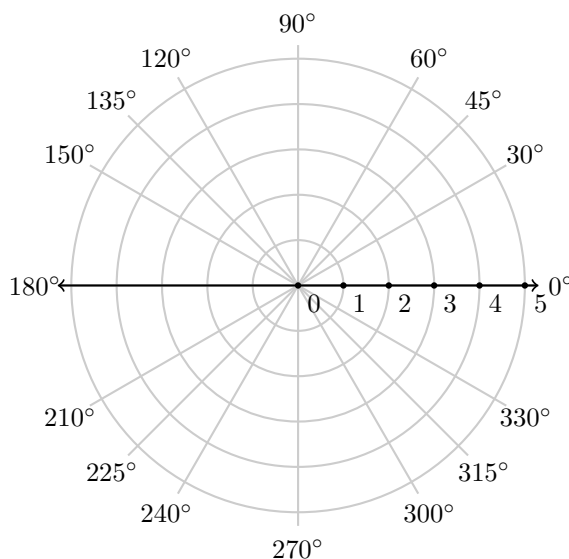
radius, but in Cartesian coordinates it can be difficult to work with since every point on the (unit) circle takes the form  $(x, \pm\sqrt{1-x^2})$ . This square root makes integrals and derivatives complicated. The main problem lies in the fact that we are trying to describe a circle using a rectangular coordinate system. So our goal in this section is to find a more suitable coordinate system for circles.

A circle can be described as the set of all points some fixed distance  $r$  from a given point, and we can specify any point on the circle by using an angle  $\theta$ :



**Figure 8.3.1:** Using distance and angle to specify a point.

Any point  $P$  on the circle can be described solely using the distance  $r$  and the direction  $\theta$ . This leads directly to the idea of polar coordinates. In polar coordinates, the  $xy$ -plane is replaced with an  $r\theta$ -plane, or **polar plane**. Each point in the polar plane has the form  $(r, \theta)$ , where the **radial coordinate**  $r$  denotes the distance from the origin, or **pole**, and the **angular coordinate**  $\theta$  determines the angle the point makes with the horizontal **polar axis**. The polar axis replaces the positive  $x$ -axis from Cartesian coordinates:



**Figure 8.3.2:** The polar plane.

In polar coordinates, positive  $\theta$  correspond to counterclockwise direction, and negative  $\theta$  correspond to clockwise direction. We also allow for  $r$  to be negative: this just means go in the direction opposite of  $\theta$ .

**Example 8.3.3 Plotting polar coordinates.** Suppose we want to plot the points  $(3, \pi)$ ,  $(-2, \frac{\pi}{4})$  and  $(2, \frac{5\pi}{4})$ . We can do so in the polar plane by remem-

bering that the first coordinate is distance from the pole, and the second coordinate is direction from the polar axis. If we plot these points, we see that  $(-2, \frac{\pi}{4})$  and  $(2, \frac{5\pi}{4})$  actually represent the same point. This is typical of polar coordinates: every point has, in general, infinitely many representations.

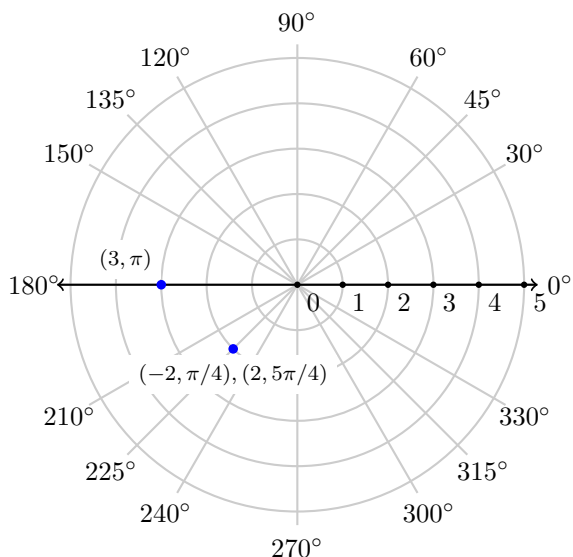
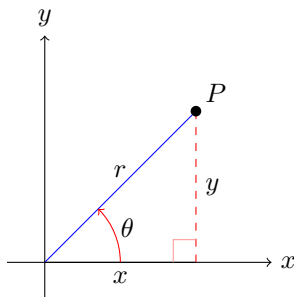


Figure 8.3.4: The plot from Example 8.3.3.

### 8.3.2 Converting Coordinates

If we want to use polar coordinates, then it'd be helpful to know how to convert between Cartesian (rectangular) coordinates and polar coordinates. The following diagram will help us to make these conversions.



So in terms of  $r, \theta$ , we have

$$x = r \cos \theta$$

$$y = r \sin \theta$$

Going in the other direction, the Pythagorean theorem tells us that  $r = \sqrt{x^2 + y^2}$ , while  $\tan \theta = \frac{y}{x}$ .

**Example 8.3.5 Converting polar to Cartesian.** Suppose we want to convert the point  $(3, \frac{\pi}{3})$  in polar coordinates to Cartesian coordinates  $(x, y)$ . Then we have

$$x = 3 \cos \frac{\pi}{3}$$

$$y = 3 \sin \frac{\pi}{3}$$

and so the Cartesian point is  $(\frac{3}{2}, \frac{3\sqrt{3}}{2})$ .

**Example 8.3.6 Converting Cartesian to polar.** Suppose we now want to convert the point  $(-\sqrt{3}, 1)$  in Cartesian coordinates to polar coordinates. Then finding  $r$  is relatively straightforward:

$$r = \sqrt{3 + 1} = 2.$$

Finding  $\theta$  requires a bit more care. We know that  $\theta$  has to satisfy  $\tan \theta = -\frac{1}{\sqrt{3}}$ . One choice that makes this work is  $\theta = -\frac{\pi}{6}$ . However, this is incorrect! Whatever  $\theta$  is needs to be consistent with the fact that our point lies in the second quadrant. So we'll pick  $\theta = \frac{5\pi}{6}$  instead. Hence one way to write the point in polar coordinates is given by  $(2, \frac{5\pi}{6})$ .

**Example 8.3.7 Converting equations.** Consider the polar equation  $r \cos \theta = 2 \sin \theta \cos \theta$ . We can convert this into a Cartesian equation using the above formulas. In particular, we use  $r \cos \theta = x$  and  $r \sin \theta = y$ . So  $\cos \theta = \frac{x}{r}$ ,  $\sin \theta = \frac{y}{r}$  and the equation becomes

$$x = \frac{2xy}{r^2} = \frac{2xy}{x^2 + y^2}.$$

### 8.3.3 Polar Curves

Now we move on to graphing polar equations.

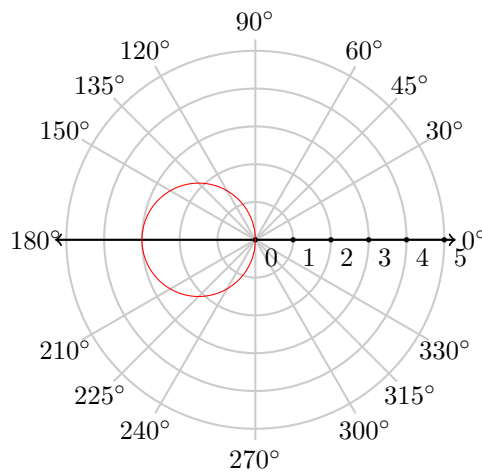
**Example 8.3.8 Graphs of constants in polar coordinates.** In Cartesian coordinates, the graphs of  $x = a$  and  $y = a$  give horizontal and vertical lines, respectively. In polar coordinates, the graphs of  $r = a$  and  $\theta = a$  have simple descriptions as well. The graph of  $r = a$  is just the set of all points  $a$  units away from the pole, so it's just a circle of radius  $|a|$  centered at the pole. Likewise, the graph of  $\theta = a$  is the set of all points that make an angle of  $a$  with the polar axis, so it's just a line through the pole.

**Example 8.3.9 Graphs from Cartesian equations.** In some cases it's beneficial to convert a polar equation to a Cartesian equation. The Cartesian equation may have a recognizable form that helps us to identify the corresponding polar graph. For example, suppose we want to describe the graph of  $r = 2a \sin \theta$  for some constant  $a$ . Then we can convert it to a Cartesian equation, in particular

$$x^2 + y^2 - 2ay = 0.$$

If we complete the square, we get  $x^2 + (y - a)^2 = a^2$ , which describes a circle centered at  $(0, a)$  (in the  $xy$ -plane) and with radius  $|a|$ . In the polar plane, this is a circle of radius  $|a|$  centered at  $(a, \frac{\pi}{2})$ . Similarly,  $r = 2a \cos \theta$  describes a circle of radius  $|a|$  centered at  $(a, 0)$ .

**Example 8.3.10 Another circle.** Suppose we want to graph  $r = -3 \cos \theta$ . From the previous example, we know that this will be a circle of radius  $\frac{3}{2}$  centered at  $(-\frac{3}{2}, 0)$ . See the following figure.

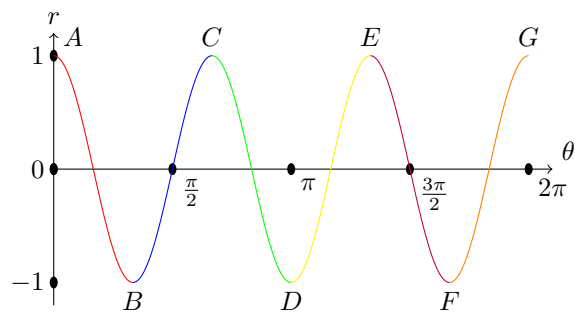


**Figure 8.3.11:** The circle from [Example 8.3.10](#).

### 8.3.4 Cartesian to Polar Method

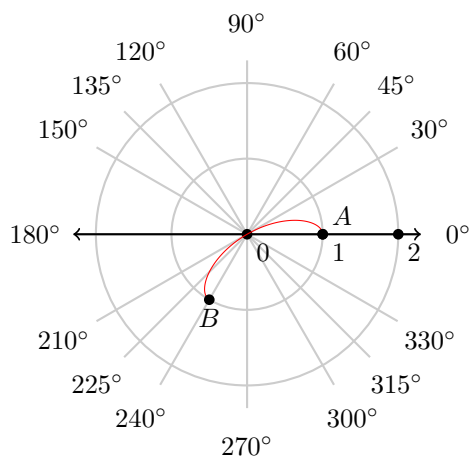
Graphing polar equations can be tricky, because it's easy to miss aspects of the graph unless you're careful. A useful method for graphing polar equations involves treating them as Cartesian equations first to get a better sense of how the graph behaves.

**Example 8.3.12 Three leaf rose.** Suppose we want to graph  $r = \cos 3\theta$ . We don't know what this looks like in the polar plane yet, but we have a pretty good idea of how it looks when treated as a Cartesian equation in the  $r\theta$ -plane (which we view as different from the polar plane!):

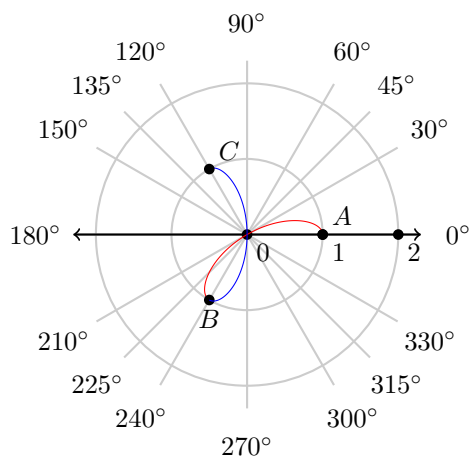


**Figure 8.3.13:** The graph of  $r = \cos 3\theta$  as a Cartesian equation.

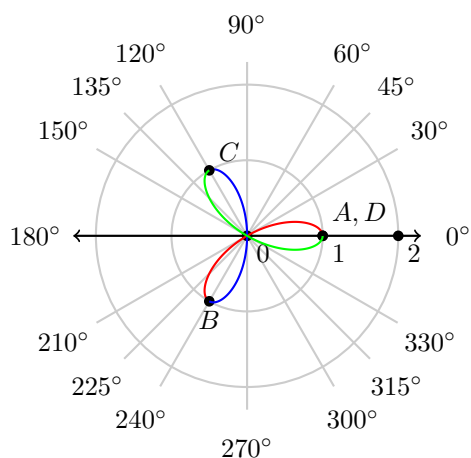
So we see that as  $\theta$  goes from 0 to  $\frac{\pi}{3}$ ,  $r$  decreases from 1 to  $-1$ , hitting 0 along the way at  $\theta = \frac{\pi}{6}$ .



Now, as  $\theta$  goes from  $\frac{\pi}{3}$  to  $\frac{2\pi}{3}$   $r$  will go from  $-1$  to  $1$ , hitting  $0$  at  $\theta = \frac{\pi}{2}$ :



Continuing in this manner lets us complete the graph:



**Figure 8.3.14:** The graph of  $r = \cos 3\theta$  in the polar plane.

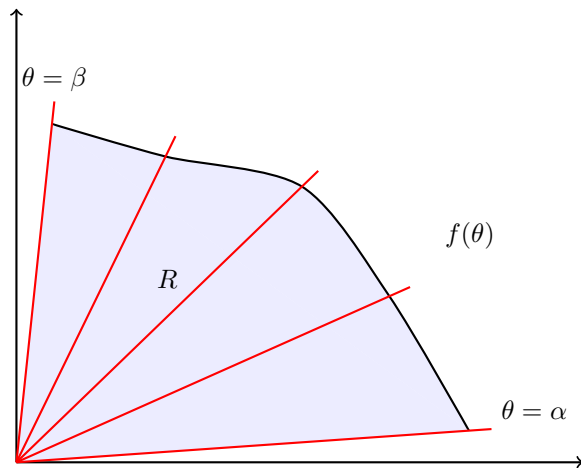
SUGGESTED PROBLEMS: 1, 3, 11, 23, 37

## 8.4 Calculus and Polar Coordinates

### 8.4.1 Areas of Polar Curves

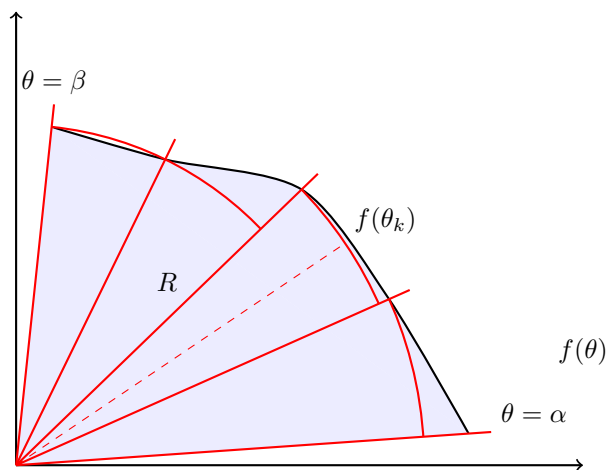
To find the area of a region swept out by a polar curve, we need to reinvent the integral so to speak. In Cartesian coordinates, the integral was based on using rectangles to approximate areas. For polar coordinates, it makes more sense to use circular sectors.

Consider a region  $R$  bounded by the polar curve  $r = f(\theta)$ :



To find the area of this region, we'll approximate it using circular sectors. Since we want to find the area between  $\theta = \alpha$  and  $\theta = \beta$ , we'll break the interval  $[\alpha, \beta]$  into  $n$  subintervals  $[\theta_i, \theta_{i+1}]$  of equal length  $\Delta\theta$ , with

$$\alpha = \theta_0 < \theta_1 < \dots < \theta_n = \beta.$$



**Figure 8.4.1:** Approximating the region with circular sectors.

The area of each slice is given by

$$\frac{\text{angle swept out by slice}}{2}(\text{radius})^2.$$

For one of our slices, this takes the form

$$\frac{\Delta\theta}{2} f(\theta_i)^2.$$

Adding these areas up gives us the approximate area of the region. So we can say that

$$\text{area of } R \approx \sum_{i=0}^n \frac{1}{2} f(\theta_i)^2 \Delta\theta.$$

As  $\Delta\theta$  gets smaller, this approximation should become exact. So we can say

$$\text{area of } R = \lim_{\Delta\theta \rightarrow 0} \sum_{i=0}^n \frac{1}{2} f(\theta_i)^2 \Delta\theta.$$

This limit is exactly the definition of an integral! So we have

$$\text{area of } R = \int_{\alpha}^{\beta} \frac{1}{2} f(\theta)^2 d\theta.$$

We can summarize this in the following theorem.

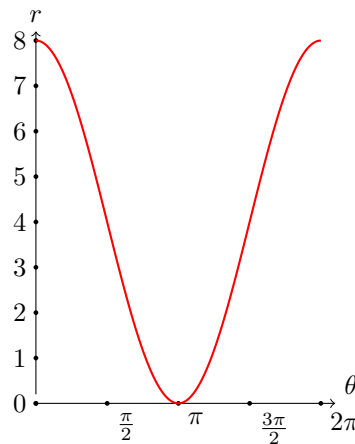
**Theorem 8.4.2 Area in Polar Coordinates.** *Suppose that  $R$  is a region bounded by the polar curves  $r = f(\theta)$  and  $r = g(\theta)$  between  $\theta = \alpha$  and  $\theta = \beta$ . If  $f$  and  $g$  are continuous and  $f(\theta) \geq g(\theta)$  for  $\alpha \leq \theta \leq \beta$ , then the area of  $R$  is given by*

$$\int_{\alpha}^{\beta} \frac{1}{2} [f(\theta)^2 - g(\theta)^2] d\theta.$$

**Example 8.4.3 Area bounded by a cardioid.** Consider the cardioid<sup>1</sup>  $r = 4 + 4 \cos \theta$ . The area is given by

$$\int_{\alpha}^{\beta} \frac{1}{2} (4 + 4 \cos \theta)^2 d\theta = \int_{\alpha}^{\beta} 8(1 + 2 \cos \theta + \cos^2 \theta) d\theta$$

To find this, we'll need to figure out the limits of integration. For polar plots, the best way to do this is by graphing. We'll graph  $r = 4 + 4 \cos \theta$  by using the Cartesian to polar methods described in [Subsection 8.3.4](#). First, we treat this as a Cartesian curve in the  $r\theta$ -plane and graph it to get

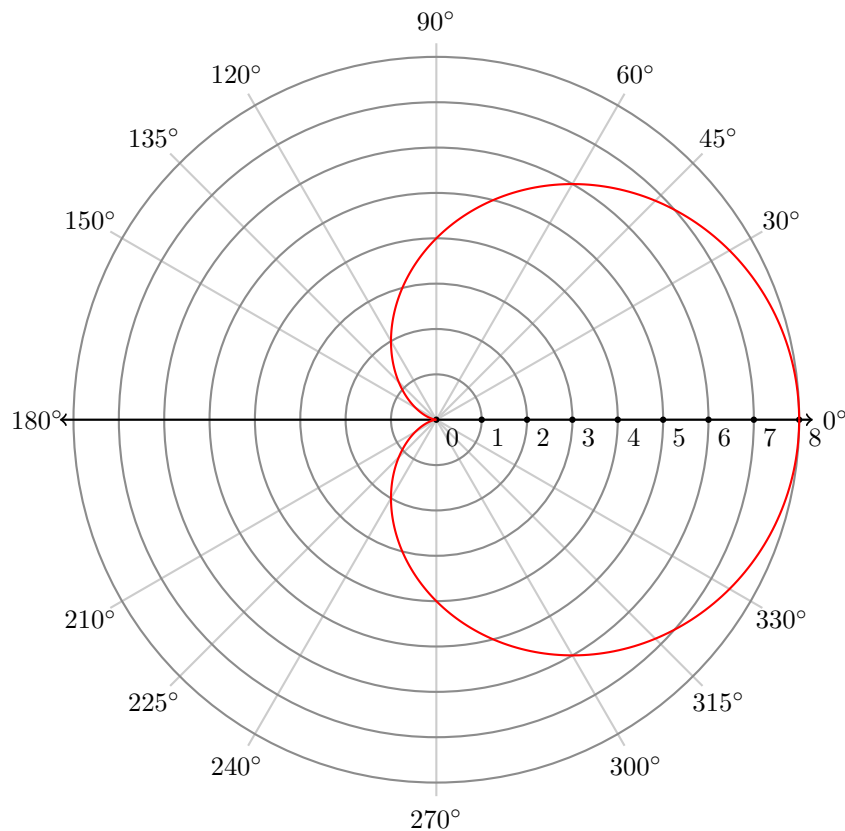


Now we can plot the corresponding polar curve:

---

<sup>1</sup>The name will become clearer once we graph this.

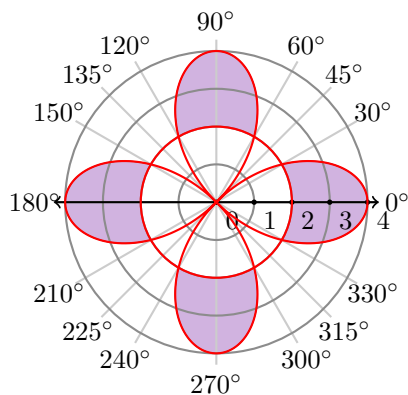




The important thing to notice is that we trace out the curve exactly once as  $\theta$  goes from 0 to  $2\pi$ , so our limits should be 0 and  $2\pi$ . Hence the area of this region is given by

$$\begin{aligned} \int_0^{2\pi} 8(1 + 2\cos\theta + \cos^2\theta) d\theta &= 8\theta \Big|_0^{2\pi} + 16\sin\theta \Big|_0^{2\pi} + 4 \int_0^{2\pi} (1 + \cos 2\theta) d\theta \\ &= 16\pi + 8\pi \\ &= 24\pi. \end{aligned}$$

**Example 8.4.4 Area inside a rose and outside a circle.** As another example of the use of [Theorem 8.4.2](#), we'll find the area of the region contained by the polar curves  $r = 4\cos 2\theta$  and outside the circle  $r = 2$ . Again, a great way to start this problem is to graph the region under consideration.



With a little bit of work, we can see that the area of one of these shaded regions is given by

$$\int_{-\frac{\pi}{6}}^{\frac{\pi}{6}} \frac{1}{2} [(4 \cos 2\theta)^2 - 4] d\theta$$

and so the area of the entire region is

$$2 \int_{-\frac{\pi}{6}}^{\frac{\pi}{6}} [(4 \cos 2\theta)^2 - 4] d\theta.$$

### 8.4.2 Arc Length in Polar Coordinates

Our goal now is to find the length of a polar curve  $r = f(\theta)$ . From [Subsection 9.8.1](#) we know how to find the length of a parametric curve, and a polar curve is really just a parametric curve with parametric equations

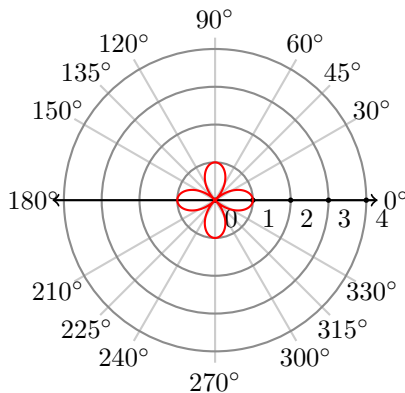
$$x = r \cos \theta = f(\theta) \cos \theta \quad \text{and} \quad y = r \sin \theta = f(\theta) \sin \theta.$$

Using the formula  $L = \int_a^b \sqrt{\left(\frac{dx}{d\theta}\right)^2 + \left(\frac{dy}{d\theta}\right)^2} d\theta$  and simplifying, we get the following result.

**Theorem 8.4.5 Arc Length of a Polar Curve.** *The length  $L$  of a polar curve  $r = f(\theta)$  from  $\theta = a$  to  $\theta = b$  is given by*

$$L = \int_a^b \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2} d\theta.$$

**Example 8.4.6 Arc length of a leaf.** Let  $r = \cos 2\theta$ . Then this gives a four-leaf rose in the polar plane. We want to find the length of one leaf of this rose. As usual, it's a good idea to graph this curve first.



**Figure 8.4.7:** The four leaf rose  $r = \cos 2\theta$ .

One leaf of the rose should be captured as  $\theta$  varies from  $-\frac{\pi}{4}$  to  $\frac{\pi}{4}$ , so using [Theorem 8.4.5](#) we can find the length of this leaf:

$$L = \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \sqrt{\cos^2 2\theta + 4 \sin^2 2\theta} d\theta = \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \sqrt{1 + 3 \sin^2 2\theta} d\theta.$$

SUGGESTED PROBLEMS: 1, 3, 7, 9, 15, 19, 35

## Chapter 9

# Vectors and Euclidean Space

Our main goal now is to extend the concepts of calculus into multiple dimensions. We start by considering the three-dimensional space  $\mathbb{R}^3$ .

### 9.1 Coordinates in 3-Space

#### 9.1.1 *xyz*-coordinates

We're used to Cartesian coordinates in the plane: each point in the plane can be represented by an  $x$ -coordinate and a  $y$ -coordinate. This representation is determined by the **coordinate axes** (the  $x$ - and  $y$ -axes). There's nothing preventing us from doing the same for three dimensions. We'll just need three coordinate axes: the  $x$ -,  $y$ - and  $z$ -axes. We typically view the  $xy$ -plane as horizontal, and the  $z$ -axis as vertical, but there's no mathematical preference either way.

Just as the  $x$ - and  $y$ -axes determined the  $xy$ -plane, we can also get the  $xz$ - and  $yz$ -planes. These planes divide space into eight **octants**. And just as any point in the  $xy$ -plane can be represented by measuring along the  $x$ -axis and  $y$ -axis to get a point  $(x, y)$ , we can do the same in space to get a point  $(x, y, z)$ . The set of all points like this is denoted by  $\mathbb{R}^3$ , and is the **three-dimensional coordinate system**.

Given any point  $(a, b, c)$  in  $\mathbb{R}^3$ , we can find its **projections** onto any of the coordinate axes or coordinate planes without too much trouble. The projection onto a line or plane is the point on that line or plane that is closest to the original point  $(a, b, c)$ .

**Example 9.1.1 Projection.** Find the projection of the point  $(-1, 3, 4)$  onto the  $yz$ -plane.

**Solution.** Here's how we can find the projection of the point  $(-1, 3, 4)$  onto the  $yz$ -plane. Note that the  $yz$ -plane is just the set of all points with  $x$ -coordinate equal to 0, so the projection of  $(-1, 3, 4)$  onto the  $yz$ -plane is the point  $(0, 3, 4)$ .

**Example 9.1.2 Equations in space.** Sketch  $y = 1$  in  $\mathbb{R}^3$ .

**Solution.** This is just the set of all points in  $\mathbb{R}^3$  of the form  $(x, 1, z)$ . This forms a **plane** in  $\mathbb{R}^3$ .

```
x, z = var('x, z')
parametric_plot3d((x, 1, z), (x, -4, 4), (z, -4, 4))
```

**Example 9.1.3 More equations in space.** Sketch  $x + y = 2$  in  $\mathbb{R}^3$ .

**Solution.** We can do so as follows. First, sketch  $x + y = 2$  in the  $xy$ -plane, which will just be the line  $y = 2 - x$ . Then the surface in  $\mathbb{R}^3$  represented by  $x + y = 2$  is actually the surface consisting of all points directly above and directly below the line  $y = 2 - x$  in the  $xy$ -plane.

**Example 9.1.4 Intersection of a sphere and a plane.** Describe the intersection of the sphere given by the equation  $(x + 2)^2 + (y - 3)^2 + z^2 = 9$  with the  $xz$ -plane.

**Solution.** We can do this without too much trouble if we remember that the  $xz$ -plane is just the set of all points with  $y$ -coordinate equal to 0. So the intersection of this sphere with the  $xz$ -plane traces out the curve  $(x + 2)^2 + (0 - 3)^2 + z^2 = 9$  in the  $xz$ -plane, which is a circle of radius  $\sqrt{18}$  centered at the point  $(-2, 0, 0)$ .

## 9.1.2 The Distance Formula

Recall that the distance between two points  $P_1(x_1, y_1)$  and  $P_2(x_2, y_2)$  in  $\mathbb{R}^2$  (the  $xy$ -plane) is given by

$$|P_1P_2| = d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

This is proved using the Pythagorean theorem. We can do the same exact thing in  $\mathbb{R}^3$ !

**Theorem 9.1.5 Distances in Space.** Let  $P_1(x_1, y_1, z_1)$  and  $P_2(x_2, y_2, z_2)$  be two points in  $\mathbb{R}^3$ . Then the distance between these two points,  $|P_1P_2| = d(P_1, P_2)$ , is given by

$$|P_1P_2| = d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}.$$

**Example 9.1.6 Computing distances.** Theorem 9.1.5 lets us find the distance between the points  $P(1, 4, -2)$  and  $Q(-1, 0, 2)$  as follows:

$$d(P, Q) = \sqrt{4 + 16 + 16} = 6.$$

One important use of the distance formula in  $\mathbb{R}^3$  is that it lets us find equations of spheres. The equation of a sphere of radius  $r$  and center  $C(h, k, l)$  is given by

$$\sqrt{(x - h)^2 + (y - k)^2 + (z - l)^2} = r,$$

which is more commonly written as

$$(x - h)^2 + (y - k)^2 + (z - l)^2 = r^2.$$

**Example 9.1.7 Equation of a sphere.** The equation  $2x^2 + 2y^2 + 2z^2 = 8x - 24z + 4$  represents a sphere in  $\mathbb{R}^3$ . To see how, we can rearrange the equation and complete the square to get

$$\begin{aligned} 2x^2 + 2y^2 + 2z^2 - 8x + 24z &= 4 \Rightarrow x^2 + y^2 + z^2 - 4x + 12z = 2 \\ &\Rightarrow (x - 2)^2 + y^2 + (z - 6)^2 = 2 + 4 + 36 \\ &\Rightarrow (x - 2)^2 + y^2 + (z - 6)^2 = 42. \end{aligned}$$

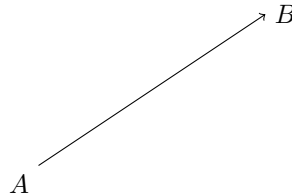
So this equation describes a sphere of radius  $\sqrt{42}$  centered at  $(2, 0, 6)$ .

**Example 9.1.8 Spherical shells.** We can also use inequalities to describe regions in addition to equalities. For example,  $3 \leq x^2 + y^2 + z^2 \leq 16$  describes the region contained between the sphere of radius  $\sqrt{3}$  and the sphere of radius 4, both centered at the origin.

SUGGESTED PROBLEMS: 1, 3, 11, 13

## 9.2 Vectors

One of our goals in this chapter is to adequately describe motion in space. A useful way to do this uses the concept of **vector**, which we think of as a quantity that has both direction and magnitude/length. A simple example would be velocity: velocity in space has a direction and also a magnitude (speed). We will typically denote vectors by using boldface letters such as  $\mathbf{x}$  or letters with a bar overhead such as  $\bar{x}$ . We represent vectors as arrows with an **initial point** and a **terminal point**:



**Figure 9.2.1:** A vector.

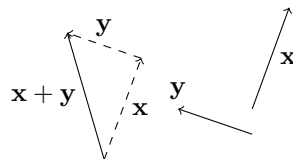
We say that two vectors  $\mathbf{x}, \mathbf{y}$  are **equivalent**, or **equal**, if they have the same magnitude and direction. We write this as  $\mathbf{x} = \mathbf{y}$ .

### 9.2.1 Addition and Scalar Multiplication

Given two vectors  $\mathbf{x} = \overrightarrow{AB}, \mathbf{y} = \overrightarrow{BC}$ , we can add them to get the new vector

$$\mathbf{x} + \mathbf{y} = \overrightarrow{AC}.$$

So the vector  $\mathbf{x} + \mathbf{y}$  is obtained by moving the tail of  $\mathbf{y}$  to the tip of  $\mathbf{x}$  and then drawing a vector from the tail of  $\mathbf{x}$  to the tip of  $\mathbf{y}$ .



**Figure 9.2.2:** Vector addition.

The sum  $\mathbf{x} + \mathbf{y}$  of vectors  $\mathbf{x}, \mathbf{y}$  can be computed using either the **triangle law**, illustrated above in Figure 9.2.2, or the similar **parallelogram law**. We can also scale vectors using **scalar multiplication**: if  $\alpha$  is a scalar<sup>1</sup> then  $\alpha\mathbf{x}$  is defined to be the vector that has the same direction as  $\mathbf{x}$  if  $\alpha > 0$  and the opposite direction if  $\alpha < 0$ , but the magnitude is rescaled by the factor  $\alpha$ .

**Example 9.2.3 Vector subtraction.** Using the previous graph, we can compute  $2\mathbf{x} - \mathbf{y}$ . We just need to scale the vectors properly and then add  $2\mathbf{x}$  to  $-\mathbf{y}$ .

### 9.2.2 Vector Components

Although it isn't too hard to add and scale vectors visually, it'll be beneficial to do the same algebraically. We can do so by breaking a vector down into its

<sup>1</sup>In other words, a number.

**components.** Consider a vector  $\mathbf{x}$  in  $\mathbb{R}^2$ , and suppose if we move it to the origin then the tip of the vector is at the point  $(3, -1)$ . Then the components of  $\mathbf{x}$  are 3 and  $-1$ , and we write

$$\mathbf{x} = \langle 3, -1 \rangle.$$

Note the use of brackets here, since technically we are saying that the vector  $\mathbf{x}$  is distinct from the point  $(3, -1)$ , even though they are closely related. We say that the **position vector** of a point is the vector whose components are the same as the coordinates of the point. Geometrically, the position vector of a point  $P$  is the vector with its tail at the origin and its tip at  $P$ .

So any vector in  $\mathbb{R}^2$  can be represented using components by  $\mathbf{x} = \langle x_1, x_2 \rangle$ . Similarly, any vector in  $\mathbb{R}^3$  can be represented as  $\mathbf{x} = \langle x_1, x_2, x_3 \rangle$ . Once you represent a vector in component form, addition and scalar multiplication is straightforward.

**Example 9.2.4 Vector addition with components.** Let  $\mathbf{a} = \langle -2, 1, 3 \rangle$  and  $\mathbf{b} = \langle 0, 2, 1 \rangle$ . Then

$$\mathbf{b} - 3\mathbf{a} = \langle 6, -1, -8 \rangle.$$

Finding magnitudes of vectors can also be done by applying the distance formula from [Section 9.1](#) to the components of the vectors. The magnitude of a vector  $\mathbf{x}$  is denoted by  $|\mathbf{x}|$  or  $\|\mathbf{x}\|$ . For example, the magnitude of  $\mathbf{b} - 3\mathbf{a}$  from the previous example is

$$\|\mathbf{b} - 3\mathbf{a}\| = \sqrt{36 + 1 + 64} = \sqrt{101}.$$

Given a vector  $\mathbf{x} = \langle x_1, x_2, x_3 \rangle$ ,

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

**Example 9.2.5 A vector equation.** We can use vectors to describe curves and surfaces. For example, let  $\mathbf{r} = \langle x, y, z \rangle$  and  $\mathbf{r}_0 = \langle a, b, c \rangle$ . Let  $r \geq 0$ . Then the set of all points  $(x, y, z)$  that satisfy the equation

$$\|\mathbf{r} - \mathbf{r}_0\| = r$$

has a very nice description: it's just the sphere of radius  $r$  centered at  $(a, b, c)$ .

**Example 9.2.6 Finding components of vectors.** Consider the points  $A(4, 0, -2)$  and  $B(4, 2, 1)$ . We want to find the components of the vector  $\overrightarrow{AB}$ . We can do this by translating  $\overrightarrow{AB}$  to the origin, which is done by subtracting from each coordinate of  $B$  the corresponding coordinate of  $A$ . So the vector  $\overrightarrow{AB}$  is given by

$$\overrightarrow{AB} = \langle 4 - 4, 2 - 0, 1 + 2 \rangle = \langle 0, 2, 3 \rangle.$$

In general, given  $A(x_1, y_1, z_1)$  and  $B(x_2, y_2, z_2)$ , the vector  $\mathbf{x} = \overrightarrow{AB}$  is given by

$$\mathbf{x} = \langle x_2 - x_1, y_2 - y_1, z_2 - z_1 \rangle.$$

**Theorem 9.2.7 Properties of Vector Addition and Scalar Multiplication.** Let  $\mathbf{x}, \mathbf{y}, \mathbf{z}$  be vectors and let  $\alpha, \beta$  be scalars. Then the following are true:

1.  $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ .
2.  $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$ .
3.  $\mathbf{x} + \mathbf{0} = \mathbf{x}$ .
4.  $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$ .
5.  $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$ .
6.  $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$ .
7.  $(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x})$ .
8.  $1\mathbf{x} = \mathbf{x}$ .

### 9.2.3 Basis Vectors and Unit Vectors

Every vector in  $\mathbb{R}^3$  can be written using three components:  $\mathbf{x} = \langle x_1, x_2, x_3 \rangle$ . Each component corresponds to a coordinate axis, and we can rewrite  $\mathbf{x}$  as a **linear combination** of three different vectors, with each vector corresponding to a coordinate axis:

$$\mathbf{x} = x_1 \langle 1, 0, 0 \rangle + x_2 \langle 0, 1, 0 \rangle + x_3 \langle 0, 0, 1 \rangle.$$

#### standard basis vectors

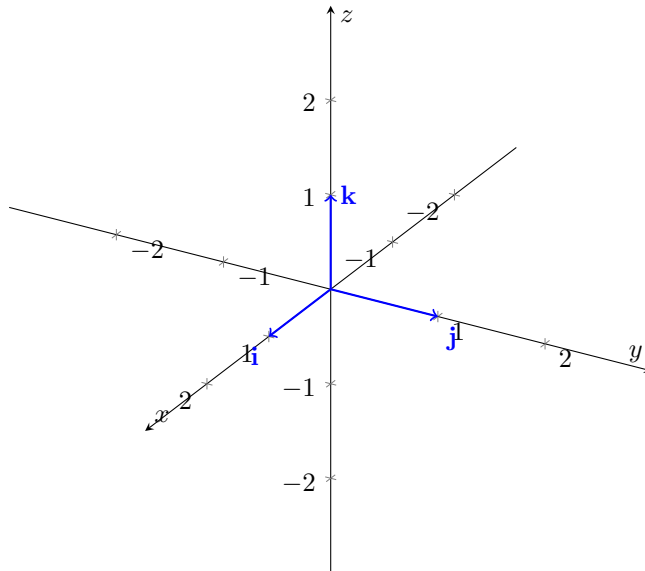
**Definition 9.2.8 Standard Basis Vectors.** The standard basis for  $\mathbb{R}^3$  is the set  $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ , where

$$\mathbf{i} = \langle 1, 0, 0 \rangle, \quad \mathbf{j} = \langle 0, 1, 0 \rangle \quad \text{and} \quad \mathbf{k} = \langle 0, 0, 1 \rangle.$$

As we've seen, every vector  $\mathbf{x}$  in  $\mathbb{R}^3$  can be expressed using only these three vectors. The standard basis has two important properties: it is perpendicular (also called **orthogonal**) and every vector in the collection has magnitude 1. In other words,

$$\|\mathbf{i}\| = \|\mathbf{j}\| = \|\mathbf{k}\| = 1.$$

These vectors are essentially designed to capture the "coordinate directions", and are plotted below.



**Figure 9.2.9:** The standard basis.

This also leads us to our next definition.

**Definition 9.2.10 Unit Vectors.** A vector  $\mathbf{x}$  is a unit vector if  $\|\mathbf{x}\| = 1$ .

Unit vectors are useful if we just need to indicate a direction, and we don't care about magnitude. Every nonzero vector can be rescaled to a unit vector: just divide the vector by its norm.

**Example 9.2.11 Direction from one point to another.** Consider the points  $A(1, 2, -1)$  and  $B(0, -3, 2)$ . Then we can find the unit vector indicating the direction from  $A$  to  $B$ . First, set

$$\mathbf{x} = \overrightarrow{AB} = \langle -1, -5, 3 \rangle.$$

Then the unit vector that gives the direction from  $A$  to  $B$  is given by

$$\begin{aligned} \mathbf{u} &= \frac{\mathbf{x}}{\|\mathbf{x}\|} \\ &= \frac{\langle -1, -5, 3 \rangle}{\sqrt{1 + 25 + 9}} \\ &= \left\langle -\frac{1}{\sqrt{37}}, -\frac{5}{\sqrt{37}}, \frac{3}{\sqrt{37}} \right\rangle. \end{aligned}$$

**Example 9.2.12 A vector equation for the unit sphere.** Using the concept of a unit vector, we can very easily describe the unit sphere<sup>2</sup> using a vector equation. If we set  $\mathbf{r} = (x, y, z)$ , then the unit sphere is just the set of all solutions of

$$\|\mathbf{r}\| = 1.$$

## 9.2.4 Applications

Many physical quantities have both a direction and a magnitude, like velocity, acceleration and forces. Vectors are ideally suited to measure these quantities.

**Example 9.2.13 Weight of a chain.** A still chain is fixed to two ends of a level divide. The tension of the chain at each fixed end can be represented by vectors pointing away from the chain. Call these tension forces  $\mathbf{T}_1$  and  $\mathbf{T}_2$ . Suppose we know that each force vector makes an angle of  $45^\circ$  with the ground on either side of the chain's fixed ends, and that the magnitude of each tension is 43 N. Then we can use vector addition to find the weight of the chain.

Let  $\mathbf{w}$  denote the weight of the chain considered as a vector (so that it's pointing down). Since the chain is still, its **resultant**<sup>3</sup> must be  $\mathbf{0}$ . So we can say that

$$\mathbf{T}_1 + \mathbf{T}_2 + \mathbf{w} = \mathbf{0}$$

or in other words

$$\mathbf{w} = -\mathbf{T}_1 - \mathbf{T}_2.$$

What we need to find is  $\|\mathbf{w}\|$ , which we can do without too much trouble if we can rewrite the tension vectors in component form. In fact, we have

$$\begin{aligned} \mathbf{T}_1 &= \langle -\|\mathbf{T}_1\| \cos 45^\circ, \|\mathbf{T}_1\| \sin 45^\circ \rangle \\ \mathbf{T}_2 &= \langle \|\mathbf{T}_2\| \cos 45^\circ, \|\mathbf{T}_2\| \sin 45^\circ \rangle \end{aligned}$$

So it follows that

$$\begin{aligned} \|\mathbf{w}\| &= \|-\mathbf{T}_1 - \mathbf{T}_2\| \\ &= \left\| \left\langle 0, -2 \cdot 43 \frac{\sqrt{2}}{2} \right\rangle \right\| \end{aligned}$$

<sup>2</sup>The sphere of radius 1 centered at the origin.

<sup>3</sup>The sum of all forces acting on the chain.



$$= 43\sqrt{2}$$

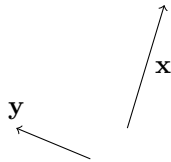
Therefore the chain weighs  $43\sqrt{2}\text{N}$ .

SUGGESTED PROBLEMS: 1–17 odd, 23, 25, 27

## 9.3 The Dot Product

### 9.3.1 Definition and Properties of the Dot Product

Suppose we're given two vectors. What we'd like to do is to come up with a measure of how much these vectors overlap. Such a measure may be useful for determining if forces are too close together, for example. So consider two vectors  $\mathbf{x} = \langle x_1, y_1, z_1 \rangle$  and  $\mathbf{y} = \langle x_2, y_2, z_2 \rangle$  as in the following diagram.



One way we can measure how much  $\mathbf{x}$  and  $\mathbf{y}$  overlap is to find  $|\mathbf{x} + \mathbf{y}|$ , or equivalently  $|\mathbf{x} + \mathbf{y}|^2$ , since this is larger if  $\mathbf{x}$  and  $\mathbf{y}$  are more closely aligned.<sup>1</sup> From [Subsection 9.2.2](#), we know that

$$|\mathbf{x} + \mathbf{y}|^2 = (x_1 + x_2)^2 + (y_1 + y_2)^2 + (z_1 + z_2)^2.$$

This simplifies out to

$$\begin{aligned} |\mathbf{x} + \mathbf{y}|^2 &= (x_1 + x_2)^2 + (y_1 + y_2)^2 + (z_1 + z_2)^2 \\ &= (x_1^2 + y_1^2 + z_1^2) + 2(x_1x_2 + y_1y_2 + z_1z_2) + (x_2^2 + y_2^2 + z_2^2) \\ &= |\mathbf{x}|^2 + 2(x_1x_2 + y_1y_2 + z_1z_2) + |\mathbf{y}|^2. \end{aligned}$$

The only part of this that could possibly depend on how closely  $\mathbf{x}$  and  $\mathbf{y}$  overlap is the middle term  $2(x_1x_2 + y_1y_2 + z_1z_2)$ . So we'll (optimistically... for now) take what's inside the parentheses and use it as the measure we're looking for.

**Definition 9.3.1 The Dot Product.** Let  $\mathbf{x} = \langle x_1, y_1, z_1 \rangle$  and  $\mathbf{y} = \langle x_2, y_2, z_2 \rangle$ . The **dot product** of  $\mathbf{x}$  with  $\mathbf{y}$ , denoted  $\mathbf{x} \cdot \mathbf{y}$ , is given by

$$\mathbf{x} \cdot \mathbf{y} = x_1x_2 + y_1y_2 + z_1z_2.$$

The dot product is also sometimes called the **scalar product** (since it always produces a scalar) or the **inner product** in other settings.

**Example 9.3.2 Computing dot products.** Let  $\mathbf{x} = \langle 1, 3, -1 \rangle$ ,  $\mathbf{y} = \langle -3, 1, 0 \rangle$  and  $\mathbf{z} = \langle 1, 1, 1 \rangle$ . Then

$$\mathbf{x} \cdot \mathbf{z} = 1 + 3 - 1 = 3 \quad \text{and} \quad \mathbf{x} \cdot \mathbf{y} = -3 + 3 + 0 = 0.$$

It also doesn't matter what order we take the vectors in for the dot product: we get the same answer regardless. However, it does matter that we only use two vectors. The dot product takes two vectors and gives a scalar. In other words, *it is impossible to take the dot product of more than two vectors at a time!* So quantities such as  $\mathbf{x} \cdot \mathbf{y} \cdot \mathbf{z}$  are not meaningful.

**Theorem 9.3.3 Properties of the Dot Product.** Let  $\mathbf{x}, \mathbf{y}$  and  $\mathbf{z}$  be vectors, and let  $\alpha$  be a scalar. Then the following properties hold:

<sup>1</sup>We are carefully ignoring the case where  $\mathbf{x}$  and  $\mathbf{y}$  point in opposite directions...

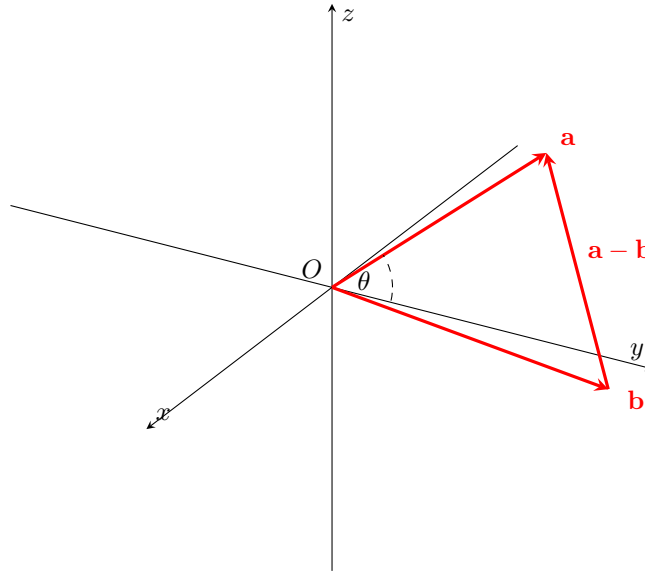
1.  $\mathbf{x} \cdot \mathbf{x} = |\mathbf{x}|^2$ .
2.  $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$ .
3.  $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$ .
4.  $(\alpha \mathbf{x}) \cdot \mathbf{y} = \alpha(\mathbf{x} \cdot \mathbf{y}) = \mathbf{x} \cdot (\alpha \mathbf{y})$ .
5.  $\mathbf{0} \cdot \mathbf{x} = 0$ .

Our goal was to define a measure for how much two given vectors align, or are correlated. The following result tells us that the dot product is actually a reasonable measure of this.

**Theorem 9.3.4 Geometry of the Dot Product.** *Let  $\mathbf{a}$  and  $\mathbf{b}$  denote vectors, and let  $0 \leq \theta \leq \pi$  denote the angle between these vectors if the tails of both are moved to the origin. Then*

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta.$$

*Proof.* Place both vectors  $\mathbf{a}$  and  $\mathbf{b}$  at the origin, and then draw the vector  $\mathbf{a} - \mathbf{b}$  from the tip of  $\mathbf{b}$  to the tip of  $\mathbf{a}$ , like so:



Then  $\mathbf{a}$ ,  $\mathbf{b}$  and  $\mathbf{a} - \mathbf{b}$  form a triangle. We want to relate the dot product  $\mathbf{a} \cdot \mathbf{b}$  with the angle  $\theta$ , so we'll start by using the Law of Cosines to get an equation involving  $\theta$ . The Law of Cosines states that

$$|\mathbf{a} - \mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2|\mathbf{a}||\mathbf{b}| \cos \theta. \quad (9.3.1)$$

Here's where the dot product comes in. Each squared magnitude can be rewritten as a dot product using [Theorem 9.3.3](#), so in particular we have

$$|\mathbf{a} - \mathbf{b}|^2 = (\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) = |\mathbf{a}|^2 - 2(\mathbf{a} \cdot \mathbf{b}) + |\mathbf{b}|^2.$$

Plugging this into (9.3.1) gives us the following:

$$|\mathbf{a}|^2 - 2(\mathbf{a} \cdot \mathbf{b}) + |\mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2|\mathbf{a}||\mathbf{b}| \cos \theta$$

Which finally simplifies down to

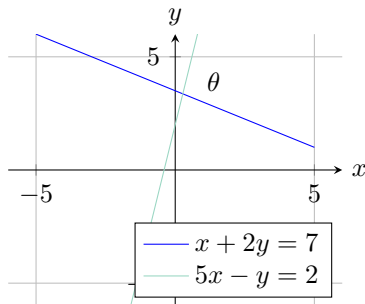
$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta.$$

□

**Remark 9.3.5.** It's usually easier to use [Definition 9.3.1](#) to compute dot products. However, [Theorem 9.3.4](#) gives us extremely useful geometric information about the dot product. For example, we get the following result very quickly: two vectors  $\mathbf{a}$  and  $\mathbf{b}$  are perpendicular if and only if  $\mathbf{a} \cdot \mathbf{b} = 0$ .

**Example 9.3.6 Checking orthogonality using the dot product.** Let  $\mathbf{a} = \langle 0, 1, -3 \rangle$  and  $\mathbf{b} = \langle 1, 1, 2 \rangle$ . Then we can say right away that these vectors are *not* orthogonal to each other since  $\mathbf{a} \cdot \mathbf{b} = -5 \neq 0$ .

**Example 9.3.7 Finding angles between lines.** Consider the lines  $x + 2y = 7$  and  $5x - y = 2$  in  $\mathbb{R}^2$ , plotted below:



Suppose we want to find the acute angle  $\theta$  these lines make with each other. We can do so by finding vectors  $\mathbf{a}$  and  $\mathbf{b}$  that point in the same directions as these lines. We'll start by finding the vector  $\mathbf{a}$  which points in the same direction as the line  $x + 2y = 7$ . To do so, we need two points on this line:  $A_1 = (7, 0)$  and  $A_2 = (5, 1)$ . So we can take  $\mathbf{a}$  to be

$$\mathbf{a} = \overrightarrow{A_1 A_2} = \langle -2, 1 \rangle.$$

Similarly, since  $B_1 = (0, -2)$  and  $B_2 = (1, 3)$  both lie on  $5x - y = 2$ , we can take

$$\mathbf{b} = \overrightarrow{B_1 B_2} = \langle 1, 5 \rangle.$$

Then by [Theorem 9.3.4](#) we know that

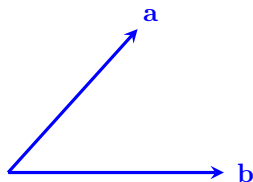
$$\cos \theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| |\mathbf{b}|} = \frac{3}{\sqrt{5} \sqrt{26}} = \frac{3}{\sqrt{130}}.$$

So the acute angle between these two lines is given by

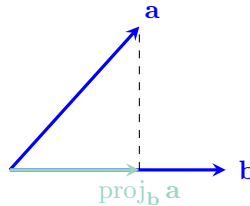
$$\theta = \cos^{-1} \frac{3}{\sqrt{130}}.$$

### 9.3.2 Projections

Consider two vectors  $\mathbf{a}$  and  $\mathbf{b}$  arranged as follows:



Now draw a line from the tip of  $\mathbf{a}$  to the point on  $\mathbf{b}$  that is closest to the tip of  $\mathbf{a}$ ; such a line must be orthogonal to  $\mathbf{b}$ . This point on  $\mathbf{b}$  defines a new vector that we call the **projection of  $\mathbf{a}$  onto  $\mathbf{b}$** ; we denote this vector by  $\text{proj}_{\mathbf{b}} \mathbf{a}$ .



The projection  $\text{proj}_{\mathbf{b}} \mathbf{a}$  can be thought of as the vector parallel to  $\mathbf{b}$  that best approximates  $\mathbf{a}$ . What we'd like to do now is to actually find a formula for this projection. If we let  $\theta$  denote the acute angle between the vectors  $\mathbf{a}$  and  $\mathbf{b}$ , then

$$|\text{proj}_{\mathbf{b}} \mathbf{a}| = |\mathbf{a}| \cos \theta.$$

Since the projection must also be parallel to  $\mathbf{b}$ , then this is enough information to state exactly what  $\text{proj}_{\mathbf{b}} \mathbf{a}$  should be:  $\text{proj}_{\mathbf{b}} \mathbf{a} = (|\mathbf{a}| \cos \theta) \frac{\mathbf{b}}{|\mathbf{b}|}$ . We can simplify this formula somewhat by making use of [Theorem 9.3.4](#).

**Theorem 9.3.8 Vector Projection Formulas.** *Let  $\mathbf{a}$  and  $\mathbf{b}$  denote vectors in  $\mathbb{R}^2$  or  $\mathbb{R}^3$ . Then the projection of  $\mathbf{a}$  onto  $\mathbf{b}$  is given by*

$$\text{proj}_{\mathbf{b}} \mathbf{a} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{b} \cdot \mathbf{b}} \mathbf{b}.$$

**Example 9.3.9 Finding vector projections.** Let  $\mathbf{a} = \langle 1, 4 \rangle$  and  $\mathbf{b} = \langle 2, 3 \rangle$ . Then by [Theorem 9.3.8](#) the projection of  $\mathbf{a}$  onto  $\mathbf{b}$  is given by

$$\text{proj}_{\mathbf{b}} \mathbf{a} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{b} \cdot \mathbf{b}} \mathbf{b} = \frac{14}{13} \langle 2, 3 \rangle.$$

Another example of vector projection is computing work done by a force. In particular, suppose we have a force  $\mathbf{F}$  and some displacement vector  $\mathbf{D}$ . We define the work done by  $\mathbf{F}$  along  $\mathbf{D}$  to be the product of the component of  $\mathbf{F}$  along  $\mathbf{D}$  times the distance moved. If we let  $\theta$  denote the acute angle between  $\mathbf{F}$  and  $\mathbf{D}$ , then the work  $W$  done is given by

$$W = (|\mathbf{F}| \cos \theta) |\mathbf{D}|,$$

which is exactly equal to  $\mathbf{F} \cdot \mathbf{D}$  by [Theorem 9.3.4](#).

**Example 9.3.10 Finding work done by a force.** Suppose a force  $\mathbf{F} = 2\mathbf{i} - 3\mathbf{j} + \mathbf{k}$  moves a particle from the point  $(2, 3, -1)$  to the point  $(1, 0, -3)$ , where the force is measured in newtons and the distance in meters. We want to find the work done by this force on the particle. To do so, we need the displacement  $\mathbf{D}$ :

$$\mathbf{D} = \langle 1 - 2, 0 - 3, -3 - (-1) \rangle = \langle -1, -3, -2 \rangle.$$

So the work done is

$$W = \mathbf{F} \cdot \mathbf{D} = -2 + 9 - 2 = 5$$

joules.

SUGGESTED PROBLEMS: 1–19 odd, 35, 39

## 9.4 The Cross Product

The dot product, and in particular [Theorem 9.3.4](#), gives us a good way to tell if two vectors are perpendicular. However, it says nothing about how to *construct* perpendicular vectors. The next vector operation, the **cross product**, is the tool we'll use for that goal.

### 9.4.1 Definition and Properties of the Cross Product

**Definition 9.4.1 The Cross Product.** Let  $\mathbf{x} = \langle x_1, y_1, z_1 \rangle$  and  $\mathbf{y} = \langle x_2, y_2, z_2 \rangle$ . Then the cross product of  $\mathbf{x}$  with  $\mathbf{y}$  is the new vector  $\mathbf{x} \times \mathbf{y}$  given by

$$\mathbf{x} \times \mathbf{y} = \langle y_1 z_2 - z_1 y_2, z_1 x_2 - x_1 z_2, x_1 y_2 - y_1 x_2 \rangle.$$

**Example 9.4.2 Cross product of basis vectors.** Let's start by computing  $\mathbf{i} \times \mathbf{k}$  using the definition. If we do so, we have

$$\mathbf{i} \times \mathbf{k} = \langle 0, -1, 0 \rangle = -\mathbf{j}.$$

On the other hand, we also have  $\mathbf{k} \times \mathbf{i} = \mathbf{j}$ . This points out the very important fact that *order matters for cross products*.

This formula is a lot to remember, so it's beneficial to find another way to express it. One way is by using **determinants**. In particular, if  $\mathbf{x} = \langle x_1, y_1, z_1 \rangle$  and  $\mathbf{y} = \langle x_2, y_2, z_2 \rangle$ , then

$$\mathbf{x} \times \mathbf{y} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix} \quad (9.4.1)$$

**Example 9.4.3 Another cross product.** [\(9.4.1\)](#) is useful to use if you're dealing with vectors that aren't as simple as the basis vectors  $\mathbf{i}, \mathbf{j}, \mathbf{k}$ . For example, let  $\mathbf{a} = \langle -1, 2, 0 \rangle$  and  $\mathbf{b} = \langle 2, 4, 1 \rangle$ . Then

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & 2 & 0 \\ 2 & 4 & 1 \end{vmatrix} \\ &= \mathbf{i} \begin{vmatrix} 2 & 0 \\ 4 & 1 \end{vmatrix} - \mathbf{j} \begin{vmatrix} -1 & 0 \\ 2 & 1 \end{vmatrix} + \mathbf{k} \begin{vmatrix} -1 & 2 \\ 2 & 4 \end{vmatrix} \\ &= 2\mathbf{i} + \mathbf{j} - 8\mathbf{k} \\ &= \langle 2, 1, -8 \rangle. \end{aligned}$$

Remember that we said the cross product is our tool for finding perpendicular vectors. So it might be nice if we made sure it actually did that. As a quick check, we'll compute  $\mathbf{a} \cdot (\mathbf{a} \times \mathbf{b})$  and  $\mathbf{b} \cdot (\mathbf{a} \times \mathbf{b})$ , with these vectors coming from [\(9.4.1\)](#). If we do so, we obtain

$$\begin{aligned} \mathbf{a} \cdot (\mathbf{a} \times \mathbf{b}) &= -2 + 2 = 0 \\ \mathbf{b} \cdot (\mathbf{a} \times \mathbf{b}) &= 4 + 4 - 8 = 0 \end{aligned}$$

Since these dot products are zero, this means that both  $\mathbf{a}$  and  $\mathbf{b}$  are perpendicular to the cross product  $\mathbf{a} \times \mathbf{b}$ . This is also true in general.

**Theorem 9.4.4 Orthogonality of the Cross Product.**  $\mathbf{a} \times \mathbf{b}$  is always orthogonal to both  $\mathbf{a}$  and  $\mathbf{b}$ .

So the cross product always produces orthogonal vectors. To determine the direction of the cross product  $\mathbf{a} \times \mathbf{b}$ , we use the **right-hand rule**: sweep your right hand from  $\mathbf{a}$  to  $\mathbf{b}$  and stick your thumb up. Then  $\mathbf{a} \times \mathbf{b}$  is parallel to your thumb.

We would also like to know the magnitude of the cross product. We can just compute it using [Definition 9.4.1](#) and find the magnitude using our usual formula. If we do so, we obtain (after a lot of simplifying!)

$$|\mathbf{a} \times \mathbf{b}|^2 = |\mathbf{a}|^2 |\mathbf{b}|^2 \sin^2 \theta,$$

which reduces to the following result.

**Theorem 9.4.5 Magnitude of the Cross Product.** *Let  $\theta$  denote the acute angle between the vectors  $\mathbf{a}$  and  $\mathbf{b}$ , so that  $0 \leq \theta \leq \pi$ . Then*

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}| |\mathbf{b}| \sin \theta.$$

So in particular, two nonzero vectors  $\mathbf{x}$  and  $\mathbf{y}$  are parallel (i.e. have  $\theta = 0$ ) if and only if  $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ .

**Example 9.4.6 Testing collinearity.** We say that three points  $P, Q$  and  $R$  are **collinear** if they all lie on the same line. Suppose we want to check if the points  $P(-2, 0, 1)$ ,  $Q(-1, 2, 4)$  and  $R(1, 6, 10)$  are collinear or not. How can we do so? If we start by defining

$$\mathbf{u} = \vec{PQ} = \langle 1, 2, 3 \rangle \quad \text{and} \quad \mathbf{v} = \vec{QR} = \langle 2, 4, 6 \rangle,$$

then we can say that all three points lie on the same line if and only if  $\mathbf{u}$  and  $\mathbf{v}$  are parallel to each other. So we'll compute their cross product to get  $\mathbf{u} \times \mathbf{v} = \mathbf{0}$ . Since these vectors are parallel, then the three given points must lie on the same line.

Another important property of the magnitude of the cross product is the following:  $|\mathbf{a} \times \mathbf{b}|$  is exactly equal to the area of the parallelogram determined by  $\mathbf{a}$  and  $\mathbf{b}$ .

**Example 9.4.7 Area of a triangle.** Suppose that we want the area of the triangle with vertices  $P(1, 2, -1)$ ,  $Q(2, 0, 0)$  and  $R(1, -1, -2)$ . To start, we need to find vectors that determine the triangle. We can use

$$\mathbf{u} = \vec{PQ} = \langle 1, -2, 1 \rangle \quad \text{and} \quad \mathbf{v} = \vec{PR} = \langle 0, -3, -1 \rangle.$$

Now, the triangle determined by  $\mathbf{u}$  and  $\mathbf{v}$  is precisely half of the parallelogram determined by these same vectors, so the area of this triangle is equal to  $\frac{1}{2} |\mathbf{u} \times \mathbf{v}|$ . We can use Sage as in the cell below to find the cross product of these vectors. Doing so, we get  $\mathbf{u} \times \mathbf{v} = \langle 5, 1, -3 \rangle$ . So the area of the triangle with vertices  $P, Q$  and  $R$  is

$$\begin{aligned} \text{area } \triangle PQR &= \frac{1}{2} |\mathbf{u} \times \mathbf{v}| \\ &= \frac{1}{2} |\langle 5, 1, -3 \rangle| \\ &= \frac{1}{2} \sqrt{35}. \end{aligned}$$

```
u = vector([1, -2, 1])
v = vector([0, -3, -1])
u.cross_product(v)
```

$(5, 1, -3)$

**Theorem 9.4.8 Properties of the Cross Product.** *Let  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  be vectors and  $c$  a scalar. Then the following properties are true:*

1.  $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$
2.  $(c\mathbf{a}) \times \mathbf{b} = c(\mathbf{a} \times \mathbf{b}) = \mathbf{a} \times (c\mathbf{b})$
3.  $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$
4.  $(\mathbf{a} + \mathbf{b}) \times \mathbf{c} = \mathbf{a} \times \mathbf{c} + \mathbf{b} \times \mathbf{c}$
5.  $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$

The last item above is an important relationship between the cross product and dot product called the **scalar triple product**. There is an important geometric significance to this new product.

**Theorem 9.4.9 Geometry of the Scalar Triple Product.** *Let  $\mathbf{a}, \mathbf{b}$  and  $\mathbf{c}$  be vectors in  $\mathbb{R}^3$ . Then  $|\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})|$  is equal to the volume of the parallelepiped determined by  $\mathbf{a}, \mathbf{b}$  and  $\mathbf{c}$ .*

**Example 9.4.10 Testing if vectors are coplanar.** We say that three vectors  $\mathbf{u}, \mathbf{v}$  and  $\mathbf{w}$  are **coplanar** if they can all lie in a single plane. For example,  $\mathbf{i}, \mathbf{k}$  and  $\langle -4, 0, 56 \rangle$  are coplanar since they lie in the  $xz$ -plane, but  $\mathbf{i}, \mathbf{j}$  and  $\mathbf{k}$  are not coplanar. Suppose we're given  $\mathbf{u} = \langle -3, 2, 1 \rangle, \mathbf{v} = \langle 1, 2, 0 \rangle$  and  $\mathbf{w} = \langle 1, 1, 1 \rangle$ . These vectors are coplanar if and only if the parallelepiped determined by these vectors has zero volume (i.e. is flat). Since

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = -9,$$

these vectors are *not* coplanar.

**Example 9.4.11 Another way to compute cross products.** Using [Theorem 9.4.8](#) and the facts that

$$\begin{aligned}\mathbf{i} \times \mathbf{j} &= \mathbf{k} \\ \mathbf{j} \times \mathbf{k} &= \mathbf{i} \\ \mathbf{k} \times \mathbf{i} &= \mathbf{j} \\ \mathbf{a} \times \mathbf{a} &= \mathbf{0}\end{aligned}$$

gives us another way to compute cross products that doesn't involve determinants. As an example, let  $\mathbf{u} = \langle 1, 4, -3 \rangle$  and  $\mathbf{v} = \langle 0, 2, 1 \rangle$ . Then

$$\begin{aligned}\mathbf{u} \times \mathbf{v} &= (\mathbf{i} + 4\mathbf{j} - 3\mathbf{k}) \times (2\mathbf{j} + \mathbf{k}) \\ &= 2\mathbf{i} \times \mathbf{j} + \mathbf{i} \times \mathbf{k} + 8\mathbf{j} \times \mathbf{j} + 4\mathbf{j} \times \mathbf{k} - 6\mathbf{k} \times \mathbf{j} - 3\mathbf{k} \times \mathbf{k} \\ &= 2\mathbf{k} - \mathbf{j} + 4\mathbf{i} - 6\mathbf{i} \\ &= \langle -2, -1, 2 \rangle.\end{aligned}$$

## 9.4.2 Torque

Consider a force  $\mathbf{F}$  acting on a rigid body at some position  $\mathbf{r}$ . This force applies a turning affect to the body, that we measure by **torque**.

**Definition 9.4.12 Torque.** The torque  $\boldsymbol{\tau}$  of a force  $\mathbf{F}$  acting at a position  $\mathbf{r}$  is defined to be the vector

$$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}.$$

As one example of torque, consider a wrench applied to a bolt. The force is exerted at the end of the wrench, and the torque is a vector that's parallel to the axis of rotation of the bolt. The torque is greater if the force is applied at a direction perpendicular (or nearly so) to that of the wrench, and smaller if the force is nearly parallel to the direction of the wrench. As a quick check, the torque is  $\mathbf{0}$  if the force is exactly parallel to the direction of the wrench, which makes sense: if we're pushing or pulling the wrench, the bolt won't rotate at all.

**Example 9.4.13 Torque and hex keys.** A hex key (Allen wrench) with a short arm of length 27 mm and a long arm length of 154 mm is applied to a screw, with the short arm attached to the screw. To turn the screw, a force of 0.5 N is applied to the long arm of the hex key turning the screw clockwise, and is exactly perpendicular to both the short arm and long arm of the hex key. We want to find the torque of this force on the screw.

One way we can do this is to imagine the screw sitting at the origin, and the hex key is (initially) in the  $xy$ -plane. Note that once the force is applied, it will begin to rotate the hex key out of the  $xy$ -plane. Now, the torque is defined by

$$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}$$

where  $\mathbf{r}$  is the vector from the screw to the point where the force  $\mathbf{F}$  is applied. We can find  $\mathbf{r}$  without too much trouble: it's  $\mathbf{r} = \langle .027, .154, 0 \rangle$ . To find the force  $\mathbf{F}$ , note that it's perpendicular to both the long arm and short arm of the screw. So a starting guess would be

$$\mathbf{F} = \langle .027, 0, 0 \rangle \times \langle 0, .154, 0 \rangle = \langle 0, 0, .042 \rangle.$$

Such a force would turn the screw clockwise, but it has the wrong magnitude. So we need to adjust it a bit:  $\mathbf{F} = \langle 0, 0, .5 \rangle$ . So the torque is given by

$$\begin{aligned} \boldsymbol{\tau} &= \mathbf{r} \times \mathbf{F} \\ &= \langle .027, .154, 0 \rangle \times \langle 0, 0, .5 \rangle \\ &= \langle .154, -.027, 0 \rangle. \end{aligned}$$

SUGGESTED PROBLEMS: 1-19 odd, 29, 41

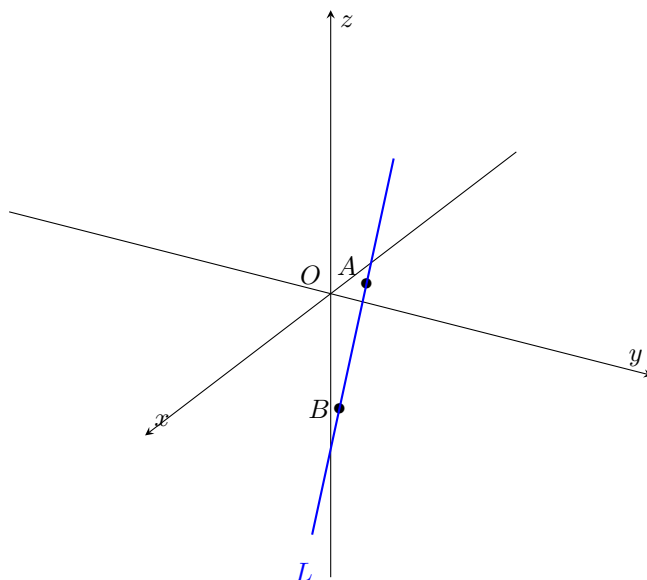
## 9.5 Equations of Lines and Planes

Recall that the equation of a line in  $\mathbb{R}^2$  has the general form  $ax + by + c = 0$ . However, this equation does not give a line in  $\mathbb{R}^3$ . In fact, it actually gives a plane, as we'll see! What we want to do now is find exactly how to describe a given line or plane.

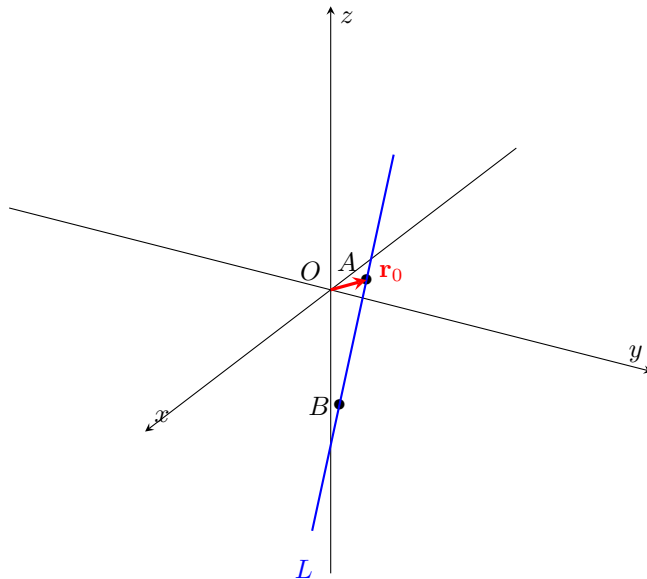
### 9.5.1 Equations of Lines

First, suppose we have a line  $L$  in  $\mathbb{R}^3$  as follows:

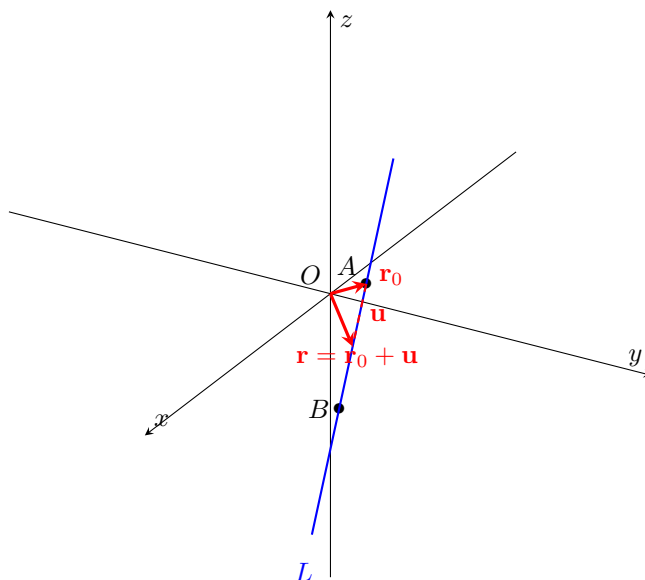




Our goal is to find an equation that describes this line. Imagine that we have a vector  $\mathbf{r}_0$  pointing from the origin to the point  $A$ , like so:



Then we can get to every other point on the line by starting from  $\mathbf{r}_0$ . In particular, if  $\mathbf{u}$  is any vector parallel to  $\vec{AB}$ , then  $\mathbf{r} = \mathbf{r}_0 + \mathbf{u}$  must also be on the line  $L$ :



So to get all possible points on  $L$ , we just need to look at all possible vectors  $\mathbf{r} = \mathbf{r}_0 + \mathbf{u}$ , where  $\mathbf{u}$  is any vector parallel to  $\vec{AB}$ . If we set  $\mathbf{v} = \vec{AB}$ ,<sup>1</sup> then

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$$

will trace out all points on  $L$  as  $t$  varies from  $-\infty$  to  $\infty$ .

**Theorem 9.5.1 Parametric Equations of a Line.** *If  $P_0(x_0, y_0, z_0)$  is a point on a line  $L$ , and if  $\mathbf{v} = \langle a, b, c \rangle$  is parallel to  $L$ , then*

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$$

*traces out  $L$ , where*

$$\mathbf{r} = \langle x, y, z \rangle, \quad \mathbf{r}_0 = \langle x_0, y_0, z_0 \rangle.$$

*We can also write this as the system of equations*

$$x = x_0 + ta, \quad y = y_0 + tb \quad \text{and} \quad z = z_0 + tc.$$

**Example 9.5.2 Equation of a line with a given direction and point.**

Suppose we want to find the equation of a line  $L$  that passes through  $(2, 5, -3)$  and is perpendicular to  $\mathbf{u} = \langle 2, 1, 2 \rangle$ . [Theorem 9.5.1](#) says that we only need to find the right choices for  $\mathbf{r}_0$  and  $\mathbf{v}$  in order to get the equation of such a line, and we can write down one of these right away:

$$\mathbf{r}_0 = \langle 2, 5, -3 \rangle.$$

So we just need to find  $\mathbf{v}$ . Since  $L$  is orthogonal to  $\mathbf{u} = \langle 2, 1, 2 \rangle$ , this means we want  $\mathbf{v}$  to be orthogonal to  $\mathbf{u}$  as well. A great way to find such a  $\mathbf{v}$  is to use the cross product. If  $\mathbf{w}$  is any nonzero vector, then  $\mathbf{u} \times \mathbf{w}$  must be orthogonal to both  $\mathbf{u}$  and  $\mathbf{w}$ , and in particular it must be orthogonal to  $\mathbf{u}$ . So we can take

$$\mathbf{v} = \langle 2, 1, 2 \rangle \times \mathbf{i} = \langle 0, 2, -1 \rangle.$$

So the equation of this line is

$$\mathbf{r} = \langle 2, 1 + 2t, 2 - t \rangle$$

As  $t$  varies, this will trace out our line.

<sup>1</sup>We don't actually have to pick  $\mathbf{v} = \vec{AB}$ , we just need  $\mathbf{v}$  to be parallel to  $\vec{AB}$ .

When we begin working with line integrals, we'll need to know how to parameterize a line segment, so we'll discuss that now. Suppose we have a line segment between two points  $P_0$  and  $P_1$ . Let  $\mathbf{r}_0$  and  $\mathbf{r}_1$  denote the corresponding position vectors of these points (i.e. the vectors whose components are given by the coordinates of these points). Then the equation of the *line* between  $P_0$  and  $P_1$  is given by

$$\mathbf{r} = \mathbf{r}_0 + t(\mathbf{r}_1 - \mathbf{r}_0) = (1 - t)\mathbf{r}_0 + t\mathbf{r}_1.$$

To get the line segment we're actually looking for, we just need to restrict  $t$  to the interval  $0 \leq t \leq 1$ . So we can say that

$$\mathbf{r} = (1 - t)\mathbf{r}_0 + t\mathbf{r}_1 \quad \text{with} \quad 0 \leq t \leq 1$$

traces out the line segment between  $P_0$  and  $P_1$ .

### 9.5.2 Equations of Planes

Now we move on to finding the equation of a plane. We'll try a vector approach to this like we did for lines, and this won't be too bad to do. We just need to keep the following in mind:

**Observation 9.5.3.** A plane in  $\mathbb{R}^3$  is uniquely determined by specifying a single point for it to pass through and a direction for it to face.

So in particular, vectors will provide a nice way to specify the direction of a plane, and any vector that does so must be orthogonal to the plane. If a (nonzero) vector  $\mathbf{n}$  is orthogonal to a given plane, then we call  $\mathbf{n}$  a **normal vector** to the plane. Now we can build up an equation of a given plane.

Let  $P_0(x_0, y_0, z_0)$  denote a point on the plane, and let  $\mathbf{n} = \langle a, b, c \rangle$  be a normal vector to the plane. If  $\mathbf{r} = \langle x, y, z \rangle$  is the position vector of any other point on the plane, then it must be true that  $\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0$ , where  $\mathbf{r}_0$  is the position vector of  $P_0$ . We can summarize this in the following theorem.

**Theorem 9.5.4 Vector and Scalar Equations of a Plane.** *If  $\mathbf{r}_0 = \langle x_0, y_0, z_0 \rangle$  is the position vector for some point on a plane with normal vector  $\mathbf{n} = \langle a, b, c \rangle$ , then the position vector  $\mathbf{r} = \langle x, y, z \rangle$  of any other point on the plane must satisfy the **vector equation of the plane**:*

$$\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0.$$

*Equivalently, any point  $(x, y, z)$  must also satisfy the equivalent **scalar equation of the plane**:*

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0.$$

**Example 9.5.5 Equation of a plane orthogonal to a line.** Consider the two lines in  $\mathbb{R}^3$  determined by the equations

$$\begin{aligned}\mathbf{r}_1 &= \langle 1, 1, 0 \rangle + t \langle 1, -1, 2 \rangle \\ \mathbf{r}_2 &= \langle 2, 0, 2 \rangle + s \langle -1, 1, 0 \rangle\end{aligned}$$

We want to find the equation of the plane that contains the intersection of these two lines and is orthogonal to the first line. By [Theorem 9.5.4](#), we need a point on the plane and a normal vector. Since the plane must contain the intersection of these two lines, that's our point. To find it, we'll set  $\mathbf{r}_1 = \mathbf{r}_2$  and go from there:

$$\mathbf{r}_1 = \mathbf{r}_2 \Rightarrow \langle -1, 1, -2 \rangle + \langle t + s, s - t, 2t \rangle = \mathbf{0}$$

$$\begin{aligned}\Rightarrow t + s - 1 &= s - t + 1 = 2t - 2 = 0 \\ \Rightarrow t &= 1 \quad \text{and} \quad s = 0\end{aligned}$$

So the position vector  $\mathbf{r}_0$  of the point where these lines intersect is given by  $\mathbf{r}_0 = \langle 2, 0, 2 \rangle$ . We also need a normal vector to the plane, but this is easy to find: since the plane needs to be orthogonal to the first line, we can just take  $\mathbf{n} = \langle 1, -1, 2 \rangle$ . So the equation of our plane is

$$\langle 1, -1, 2 \rangle \cdot \langle x - 2, y, z - 2 \rangle = 0$$

or equivalently

$$x - y + 2z - 6 = 0.$$

This plane is plotted below using the Sage cell immediately after this example.

```
from sage.manifolds.utilities import set_axes_labels
var('x,y')
P = plot3d((y-x)/2+3, (x,-3,3), (y,-3,3))
set_axes_labels(P, 'x', 'y', 'z')
```

We can rewrite the equation of a plane in the form

$$ax + by + cz + d = 0.$$

A normal vector to this plane is given by  $\mathbf{n} = \langle a, b, c \rangle$ .

**Example 9.5.6 Equation of a line orthogonal to a plane.** Suppose we're given the plane  $3x - 5z + 2 = 0$ , and we want to find the equation of the line orthogonal to this plane and passing through the point  $(3, -1, 2)$ . Then the equation of this line is

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$$

where  $\mathbf{r}_0 = \langle 3, -1, 2 \rangle$  and  $\mathbf{v}$  is a vector parallel to the line. Since the line must be orthogonal to the plane, we can take  $\mathbf{v} = \langle 3, 0, -5 \rangle$ . So the equation of this line is

$$\mathbf{r} = \langle 3, -1, 2 \rangle + \langle 3t, 0, -5t \rangle.$$

**Example 9.5.7 Equation of a plane containing a line and a point.** Suppose we're trying to find a plane containing the point  $(1, 2, 0)$  and containing the line  $\mathbf{r} = \langle 3 - t, 2, 5 + 2t \rangle$ . To get the equation of this plane, we also need a normal vector to the plane. We can find the normal vector by computing the cross product of two vectors parallel to the plane.

First, since the plane contains the line  $\langle 3 - t, 2, 5 + 2t \rangle$ , the direction of this line must be parallel to the plane. So we know that  $\mathbf{v} = \langle -1, 0, 2 \rangle$  is parallel to the plane. To get another vector parallel to this plane, we can use the given point and a second point in the plane, say  $(3, 2, 5)$ , which comes from the line (just set  $t = 0$ ). So the vector  $\mathbf{u} = \langle 3 - 1, 2 - 2, 5 - 0 \rangle = \langle 2, 0, 5 \rangle$  must also be parallel to the plane. Hence a normal vector to this plane is

$$\mathbf{n} = \mathbf{u} \times \mathbf{v} = \langle 0, 9, 0 \rangle.$$

So an equation of this plane is given by

$$9(y - 2) = 0.$$

SUGGESTED PROBLEMS: 1, 3, 5, 11, 17, 21, 27, 31

## 9.6 Cylinders and Quadric Surfaces

Before we move on to doing calculus with vectors, we'll briefly take a look at more graphs in  $\mathbb{R}^3$ . In particular, we'll look at **cylinders** and **quadric surfaces**.

### 9.6.1 Cylinders

**Definition 9.6.1 Cylinders.** A cylinder is the collection of all lines parallel to a given line and passing through some plane curve.

A basic example of a cylinder is the set of all lines passing through the unit circle in the  $xy$ -plane and parallel to the  $z$ -axis. In  $\mathbb{R}^3$ , this is just the graph of  $x^2 + y^2 = 1$ . The graph of this cylinder is provided below.

```
# Code shamelessly stolen from this stackoverflow post:
# http://stackoverflow.com/questions/26989131/add-cylinder-to-plot
import matplotlib.pyplot as plt
import numpy as np
from mpl_toolkits.mplot3d import Axes3D

plt.rc('text', usetex=True)

fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')

# Cylinder
x=np.linspace(-1, 1, 100)
z=np.linspace(-2, 2, 100)
Xc, Zc=np.meshgrid(x, z)
Yc = np.sqrt(1-Xc**2)

# Draw parameters
rstride = 20
cstride = 10
ax.plot_surface(Xc, Yc, Zc, alpha=0.2, rstride=rstride,
               cstride=cstride)
ax.plot_surface(Xc, -Yc, Zc, alpha=0.2, rstride=rstride,
               cstride=cstride)

ax.set_xlabel(r'$x$')
ax.set_ylabel(r'$y$')
ax.set_zlabel(r'$z$')
plt.show()
```

For our purposes, equations that give a cylinder will often be missing a variable.

**Example 9.6.2 A sinusoidal cylinder.** Consider the equation  $z = \sin y$  in  $\mathbb{R}^3$ . This equation is missing the variable  $x$ , which suggests that the graph of this equation should be a cylinder. However, it's not going to look like the cylinders we may be used to at this point. In fact, this is just the set of all lines passing through the curve  $z = \sin y$  in the  $yz$ -plane and parallel to the  $x$ -axis. It's graph is given below.

```
# Code shamelessly adapted from this stackoverflow post:
# http://stackoverflow.com/questions/26989131/add-cylinder-to-plot
import matplotlib.pyplot as plt
import numpy as np
```

```

from mpl_toolkits.mplot3d import Axes3D

plt.rc('text', usetex=True)

fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')

# Cylinder
x=np.linspace(-1, 1, 100)
y=np.linspace(-2, 2, 100)
Xc, Yc=np.meshgrid(x, y)
Zc = np.sin(Yc)

# Draw parameters
rstride = 20
cstride = 10
ax.plot_surface(Xc, Yc, Zc, alpha=0.2, rstride=rstride,
               cstride=cstride)

ax.set_xlabel(r'$x$')
ax.set_ylabel(r'$y$')
ax.set_zlabel(r'$z$')
plt.show()

```

**Example 9.6.3 Another cylinder.** Consider the cylinder given by the set of all lines passing through the plane curve  $y = 3x - 1$  in the  $xy$ -plane and parallel to the line in  $\mathbb{R}^3$  defined by the equation

$$\mathbf{r} = \langle 3 - t, t, 3 \rangle.$$

What does this cylinder look like? Well, we can view it as essentially a "sheet" of lines cutting through the  $xy$ -plane at the line  $y = 3x - 1$ . If we try to imagine this, then this suggests that this cylinder should probably be a plane! In fact, this cylinder is exactly the plane containing the point  $(0, -1, 0)$  and parallel to the line  $\mathbf{r}$  given above.

## 9.6.2 Quadric Surfaces

A **quadric surface** is any surface that is the graph of an equation of the form

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz + 0 = 0.$$

A useful tool for graphing quadric surfaces (and others in  $\mathbb{R}^3$ ) is the concept of a **trace**, which is what the curve looks like in a plane parallel to the one the coordinate planes. This amounts to setting either,  $x$ ,  $y$  or  $z$  equal to a constant and graphing the resulting equation.

**Example 9.6.4 An ellipsoid.** Consider the equation

$$4x^2 + \frac{y^2}{9} + \frac{z^2}{25} = 1.$$

If we want to graph this, we can graph a few of its traces to get an idea of what it looks like. Let's graph traces parallel the  $xz$ -plane to start. This means we'll set  $y$  equal to different constants. For  $y = 0$ , we get the equation

$$4x^2 + \frac{z^2}{25} = 1$$

which we rewrite as

$$\frac{x^2}{(\frac{1}{2})^2} + \frac{z^2}{25} = 1.$$

This is an ellipse in the  $xz$ -plane, with minor axis  $\frac{1}{2}$  and major axis 5. We can graph another trace, say in the  $xy$ -plane, we get

$$\frac{x^2}{(\frac{1}{2})^2} + \frac{y^2}{9} = 1,$$

which is an ellipse with minor axis  $\frac{1}{2}$  and major axis 3. Similarly, in the  $yz$ -plane we have an ellipse with minor axis 3 and major axis 5. Putting these together gives us a rough idea of the shape of this surface, which we call an **ellipsoid**.

**Example 9.6.5 Region between surfaces.** Suppose we want to sketch the region between the surface  $z = \sqrt{x^2 + y^2}$  and  $x^2 + y^2 = 1$  for  $1 \leq z \leq 2$ . First, we can graph  $z = \sqrt{x^2 + y^2}$ . If we look at the horizontal traces of this surface, we get circles of varying radii. As  $z$  increases, the radii of these circles increase as well. This surface is just a cone! So we're describing the region of this cone bounded between  $z = 1$  and  $z = 2$ , and contained inside the cylinder  $x^2 + y^2 = 1$ .

SUGGESTED PROBLEMS: 1-7 odd, 11, 29

## 9.7 Vector Functions

Recall from [Subsection 9.5.1](#) that the equation of a line can be written as

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v} = \langle x_0 + at, y_0 + bt, z_0 + ct \rangle.$$

This is our first example of a **vector function**. Vector functions are functions of the form

$$\mathbf{r}(t) = \langle f(t), g(t), h(t) \rangle,$$

and graphs of vector functions are called **space curves**. We call  $f, g, h$  the **component functions** of  $\mathbf{r}$ . We're interested in how these curves change, which means we're interested in how to do calculus on space curves. Although these curves live in  $\mathbb{R}^3$ , there's still only one independent variable:  $t$ . So much of what we learned in Calculus I applies to space curves.

### 9.7.1 Limits with Space Curves

We can take limits with vector functions just as we can with regular functions.

Let  $\mathbf{r}(t) = \langle f(t), g(t), h(t) \rangle$ . Then

$$\lim_{t \rightarrow a} \mathbf{r}(t) = \left\langle \lim_{t \rightarrow a} f(t), \lim_{t \rightarrow a} g(t), \lim_{t \rightarrow a} h(t) \right\rangle. \quad (9.7.1)$$

In other words, if you want to take the limit of a vector function you can just take the limits of the component functions.

**Example 9.7.1 Limit of a vector function.** Let

$$\mathbf{r}(t) = \frac{t^2 - t}{t - 1} \mathbf{i} + \sqrt{t + 8} \mathbf{j} - \frac{\sin \pi t}{\ln t} \mathbf{k}.$$

Suppose we want to find  $\lim_{t \rightarrow 1} \mathbf{r}(t)$ . Then we just need to take the limit of each component. So

$$\lim_{t \rightarrow 1} \mathbf{r}(t) = 2\mathbf{i} + 3\mathbf{j} - \pi\mathbf{k}.$$

Just as in Calculus I, we say that a vector function  $\mathbf{r}(t)$  is **continuous** at  $t = a$  if  $\lim_{t \rightarrow a} \mathbf{r}(t) = \mathbf{r}(a)$ . In general, a vector function is continuous wherever *all* of its components functions are continuous.

**Example 9.7.2 A horizontal helix.** Let  $\mathbf{r}(t) = \langle \sin \pi t, t, \cos \pi t \rangle$ , and suppose we want to sketch this function. One way to do so is to plug in values for  $t$  and connect the resulting points with a curve, but we can also do the following to get an idea of what this looks like. First, note that we have  $x = \sin \pi t$  and  $z = \cos \pi t$ . So  $x^2 + z^2 = 1$ , which means that this looks like the unit circle in the  $xz$ -plane, traced *clockwise*. Since we also have  $y = t$ , this curve moves farther along the  $y$ -axis as  $t$  increases. If we trace this out, we get a helix (see the below plot). We can also see from the graph that it has no jumps or gaps, so  $\mathbf{r}(t)$  is continuous everywhere.

```
# Code shamelessly adapted from this stackoverflow post:
# http://stackoverflow.com/questions/26989131/add-cylinder-to-plot
import matplotlib.pyplot as plt
import numpy as np
from mpl_toolkits.mplot3d import Axes3D

plt.rc('text', usetex=True)

fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')

# Define variables/components
t = np.linspace(-5, 5, 500)
x = np.sin(np.pi*t)
y = t
z = np.cos(np.pi*t)

# Draw helix
ax.plot(x, y, z, label=r'\mathbf{r}(t)')
ax.legend()

ax.set_xlabel(r'$x$')
ax.set_ylabel(r'$y$')
ax.set_zlabel(r'$z$')
plt.show()
```

**Example 9.7.3 Finding vector functions.** Consider the cylinder  $x^2 + y^2 = 4$  and the surface  $z = xy$ , and suppose we want to trace out their intersection with a vector function. Here's how we can do this. First, we'll come up with the  $x$  and  $y$  components of  $\mathbf{r}(t)$ . Since  $x^2 + y^2 = 4$ , this suggests that we should take

$$x = 2 \cos t$$

$$y = 2 \sin t$$

So that's two down, one to go. To get  $z$ , we just need to use the equation  $z = xy$ . So

$$z = xy = 4 \cos t \sin t.$$

So our vector function is

$$\mathbf{r}(t) = \langle 2 \cos t, 2 \sin t, 2 \sin 2t \rangle.$$

This is also plotted below.



```
# Code shamelessly adapted from this stackoverflow post:
# http://stackoverflow.com/questions/26989131/add-cylinder-to-plot
# If I change it enough it becomes MY code
import matplotlib.pyplot as plt
import numpy as np
from mpl_toolkits.mplot3d import Axes3D

plt.rc('text', usetex=True)

fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')

# Define variables/components
t = np.linspace(-5,5,500)
x = 2*np.cos(t)
y = 2*np.sin(t)
z = x*y

# Draw curve
ax.plot(x, y, z, label=r'\mathbf{r}(t)')
ax.legend()

ax.set_xlabel(r'$x$')
ax.set_ylabel(r'$y$')
ax.set_zlabel(r'$z$')
plt.show()
```

### 9.7.2 Derivatives with Space Curves

Now that we know how to take limits with vector functions, we can take derivatives as well.

**Definition 9.7.4 Derivatives of Vector Functions.** Let  $\mathbf{r}(t)$  denote a vector function. The **derivative** of  $\mathbf{r}$  is the new vector function  $\mathbf{r}'$  given by

$$\mathbf{r}'(t) = \frac{d\mathbf{r}}{dt} = \lim_{h \rightarrow 0} \frac{\mathbf{r}(t+h) - \mathbf{r}(t)}{h},$$

assuming that the limit exists. If this limit exists, we say that  $\mathbf{r}$  is **differentiable**.

If  $\mathbf{r}(t) = \langle f(t), g(t), h(t) \rangle$ , then  $\mathbf{r}$  is differentiable if and only if  $f, g, h$  are, and  $\mathbf{r}'(t) = \langle f'(t), g'(t), h'(t) \rangle$ . Just as in Calculus I, the derivative represents how quickly a space curve is changing at some value of  $t$ . However, derivatives of vector functions also carry information about the *direction* a curve is moving. We call  $\mathbf{r}'(t)$  the **tangent vector** to  $\mathbf{r}(t)$ . In particular,  $\mathbf{r}'(t)$  is parallel to the space curve  $\mathbf{r}$  at  $t$ , and its magnitude  $|\mathbf{r}'(t)|$  represents how quickly the curve is changing at  $t$ . If we only care about direction, then we can define the **unit tangent**  $\mathbf{T}(t)$ , which is given by

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|}.$$

We also have the usual ideas from Calculus I and physics regarding motion: velocity is the derivative of position and acceleration is the derivative of velocity.

**Example 9.7.5 Velocity on a saddle.** A particle moves counterclockwise along the "saddle"  $\mathbf{r}(t) = \langle 2 \cos t, 2 \sin t, 2 \sin 2t \rangle$ . We want its velocity at  $t = \frac{\pi}{2}$ . First, find  $\mathbf{r}'$  to get

$$\mathbf{r}'(t) = \langle -2 \sin t, 2 \cos t, 4 \cos 2t \rangle.$$

At  $t = \frac{\pi}{2}$ , we have the velocity vector

$$\mathbf{r}'\left(\frac{\pi}{2}\right) = \langle -2, 0, -4 \rangle.$$

So at the point  $t = \frac{\pi}{2}$ , the space curve is parallel to the vector  $\langle -2, 0, 4 \rangle$ . In other words, the particle is moving in this direction at  $t = \frac{\pi}{2}$ .

**Example 9.7.6 Tangents on a circle.** A particle moves along the circle  $y^2 + z^2 = 5$  in the  $yz$ -plane, counterclockwise and with an angular frequency of  $5\pi \frac{\text{rad}}{\text{s}}$ . Then we can assume that its position is described by

$$\mathbf{r}(t) = \langle 0, \sqrt{5} \cos 5\pi t, \sqrt{5} \sin 5\pi t \rangle.$$

Suppose we want to find the direction this particle is going at any given moment. Then we can just find the unit tangent vector  $\mathbf{T}$ :

$$\mathbf{T}(t) = \frac{\langle 0, -5\sqrt{5} \sin 5\pi t, 5\sqrt{5} \cos 5\pi t \rangle}{\sqrt{125}}.$$

We also have derivative rules for vector functions, based off of the familiar formulas from Calculus I.

**Theorem 9.7.7 Vector Derivative Rules.** *Let  $\mathbf{u}$  and  $\mathbf{v}$  be differentiable vector functions,  $c$  be a scalar and let  $f(t)$  be a differentiable (scalar) function. Then the following formulas hold:*

1.  $\frac{d}{dt}[\mathbf{u}(t) + \mathbf{v}(t)] = \mathbf{u}'(t) + \mathbf{v}'(t)$
2.  $\frac{d}{dt}[c\mathbf{u}(t)] = c\mathbf{u}'(t)$
3.  $\frac{d}{dt}[f(t)\mathbf{u}(t)] = f'(t)\mathbf{u}(t) + f(t)\mathbf{u}'(t)$
4.  $\frac{d}{dt}[\mathbf{u}(t) \cdot \mathbf{v}(t)] = \mathbf{u}'(t) \cdot \mathbf{v}(t) + \mathbf{u}(t) \cdot \mathbf{v}'(t)$
5.  $\frac{d}{dt}[\mathbf{u}(t) \times \mathbf{v}(t)] = \mathbf{u}'(t) \times \mathbf{v}(t) + \mathbf{u}(t) \times \mathbf{v}'(t)$
6.  $\frac{d}{dt}\mathbf{u}(f(t)) = f'(t)\mathbf{u}'(f(t))$

### 9.7.3 Integrals with Space Curves

We can also integrate vector functions without too much trouble. Just as taking the derivative of a vector function reduces down to differentiating each component, integrating a vector function reduces down to integrating each component. If  $\mathbf{r}(t) = \langle f(t), g(t), h(t) \rangle$ , then

$$\int_a^b \mathbf{r}(t) dt = \left\langle \int_a^b f(t) dt, \int_a^b g(t) dt, \int_a^b h(t) dt \right\rangle. \quad (9.7.2)$$

SUGGESTED PROBLEMS: 1, 3, 5, 7, 17, 21, 23, 37, 41, 45, 59, 61

## 9.8 Arc Length and Curvature

### 9.8.1 Arc Length

Suppose some particle travels along the space curve given by  $\mathbf{r}(t)$  for  $a \leq t \leq b$ . What we'd like to do is determine how far the particle travels over this time interval. Recall that distance is just speed multiplied by time, and the speed of this particle is given by  $|\mathbf{r}'(t)|$ . So we can imagine that the distance traveled by this particle over some infinitesimal time interval  $dt$  to be given by  $|\mathbf{r}'(t)| dt$ . So adding up all of these distances from  $t = a$  to  $t = b$  should give us the arc length. This suggests (but does not prove!) the following formula for arc length:

$$L = \int_a^b |\mathbf{r}'(t)| dt. \quad (9.8.1)$$

(9.8.1) gives the length  $L$  of the space curve  $\mathbf{r}(t)$  from  $t = a$  to  $t = b$ , assuming that the integral exists.

**Example 9.8.1 Arc length of a helix.** Suppose we want to find the arc length of the helix  $\mathbf{r}(t) = \langle \cos 3t, \sin 3t, t \rangle$  from  $t = 0$  to  $t = 1$ . Then this is given by

$$L = \int_0^1 |\langle -3 \sin 3t, 3 \cos 3t, 1 \rangle| dt = \int_0^1 \sqrt{10} dt = \sqrt{10}.$$

One thing we'd like to do now is to parametrize a space curve  $\mathbf{r}(t)$  with respect to arc length. Here's what we mean by this: given some  $\mathbf{r}(t)$  with  $a \leq t \leq b$ , we define its **arc length function**  $s(t)$  by

$$s(t) = \int_a^t |\mathbf{r}'(u)| du.$$

If we can then solve for  $t$  in terms of  $s$ , this parametrizes the curve  $\mathbf{r}(t)$  in terms of the arc length variable  $s$ .

**Example 9.8.2 Reparametrizing a space curve.** Suppose we're given the space curve

$$\mathbf{r}(t) = \left( \frac{2}{t^2 + 1} - 1 \right) \mathbf{i} + \frac{2t}{t^2 + 1} \mathbf{j}$$

which starts at  $(1, 0)$ , (so  $t$  starts at 0) and we want to find the point that is  $\frac{\pi}{2}$  units along the curve in the positive direction. Then we can do this by reparametrizing the curve using arc length. Here's how. First, we find the arc length function  $s(t)$ :

$$\begin{aligned} s(t) &= \int_0^t |\mathbf{r}'(u)| du \\ &= \int_0^t \sqrt{\frac{4u^2 + 4(u^2 + 1)^2 - 16u^2(u^2 + 1) + 16u^4}{(u^2 + 1)^4}} du \\ &= \int_0^t \frac{2}{u^2 + 1} du \\ &= 2 \tan^{-1} t \end{aligned}$$

Since  $s = 2 \tan^{-1} t$ , we get  $t = t(s) = \tan \frac{s}{2}$ . So

$$\mathbf{r}(t(s)) = \left( \frac{2}{\tan^2 \frac{s}{2} + 1} - 1 \right) \mathbf{i} + \frac{2 \tan \frac{s}{2}}{\tan^2 \frac{s}{2} + 1} \mathbf{j}$$

reparametrizes the space curve  $\mathbf{r}$  in terms of arc length. So the point on the curve that is  $\frac{\pi}{2}$  units along in the positive direction is given by

$$\mathbf{r}(t(\pi/2)) = \langle 0, 1 \rangle.$$

```
# Code shamelessly adapted from this stackoverflow post:
# http://stackoverflow.com/questions/26989131/add-cylinder-to-plot
# If I change it enough it becomes MY code
import matplotlib.pyplot as plt
import numpy as np
from mpl_toolkits.mplot3d import Axes3D

plt.rc('text', usetex=True)

fig = plt.figure()
ax = fig.add_subplot(111)

# Define variables/components
t = np.linspace(-5, 5, 500)
x = 2/(t^2+1) - 1
y = 2*t/(t^2+1)

# Draw curve
ax.plot(x, y, label=r'$\mathbf{r}(t)$')
ax.legend()

ax.set_xlabel(r'$x$')
ax.set_ylabel(r'$y$')
plt.show()
```

## 9.8.2 Curvature

What we want to do now is measure how much a curve "turns" at some point. First, we call a space curve  $\mathbf{r}(t)$  **smooth** on some interval  $I$  if  $\mathbf{r}'(t) \neq 0$  for any  $t$  in  $I$ . Now, if we want to measure how quickly a curve is changing direction then we can use the unit tangent vector  $\mathbf{T}(t)$  to measure this. In particular, the derivative of the unit tangent should provide a good measure of how quickly a curve turns. However, we don't want to let the specific parametrization of the curve affect this; in other words, the rate at which a curve is turning should not depend on the speed at which the curve is traveled. This leads to the following definition.

**Definition 9.8.3 Curvature.** Let  $\mathbf{r}(t)$  denote a smooth curve on some interval  $I$ . The **curvature** of  $\mathbf{r}(t)$  on  $I$  is defined to be the function  $\kappa(t)$  given by

$$\kappa(t) = \frac{|\mathbf{T}'(t)|}{|\mathbf{r}'(t)|},$$

where  $\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|}$  is the unit tangent to the curve.

**Example 9.8.4 Computing a curvature.** Consider the curve  $\mathbf{r}(t) = \langle t, \frac{1}{2}t^2, t^2 \rangle$ . We'll try to find the curvature by making use of [Definition 9.8.3](#). First, we need to find  $\mathbf{T}$  so we can take its derivative:

$$\mathbf{T}(t) = \frac{\langle 1, t, 2t \rangle}{\sqrt{1 + t^2 + 4t^2}}.$$

Therefore

$$\begin{aligned}\mathbf{T}'(t) &= -\frac{1}{2}(1+5t^2)^{-3/2}(10t)\langle 1, t, 2t \rangle + (1+5t^2)^{-1/2}\langle 0, 1, 2 \rangle \\ &= -5t(1+5t^2)^{-3/2}\langle 1, t, 2t \rangle + (1+5t^2)^{-1/2}\langle 0, 1, 2 \rangle \\ &= (1+5t^2)^{-3/2}\langle -5t, 1, 2 \rangle\end{aligned}$$

So the curvature is given by

$$\begin{aligned}\kappa(t) &= \frac{(1+5t^2)^{-3/2}\sqrt{25t^2+5}}{(1+5t^2)^{1/2}} \\ &= \sqrt{5}(1+5t^2)^{-3/2}\end{aligned}$$

**Example 9.8.5 Using curvature.** Consider the curve  $\mathbf{r}(t) = \langle t, \frac{1}{2}t^2, t^2 \rangle$  once again. We'll try to find where this curve turns the fastest. To do this, we look at the curvature  $\kappa(t)$  we found from [Example 9.8.4](#):

$$\kappa(t) = \sqrt{\frac{5}{(1+5t^2)^3}}.$$

This curve is turning the fastest precisely where the curvature is largest, and this happens at  $t = 0$ .

We can use [Definition 9.8.3](#) to compute curvatures, but we've seen that it can be pretty awful. So we'd like to use another formula if we can; thankfully, there is another option.

**Theorem 9.8.6 Alternative Curvature Formula.** *The curvature of  $\mathbf{r}(t)$  in  $\mathbb{R}^3$  is given by*

$$\kappa(t) = \frac{|\mathbf{r}'(t) \times \mathbf{r}''(t)|}{|\mathbf{r}'(t)|^3}.$$

The formula in [Theorem 9.8.6](#) may look worse than the formula given in [Definition 9.8.3](#), but it has one significant advantage: we don't need to differentiate any magnitudes, which is required in the previous formula to compute  $\mathbf{T}'(t)$ .

**Example 9.8.7 Using the alternative formula.** Let  $\mathbf{r}(t) = t\mathbf{i} + t^2\mathbf{j} + e^t\mathbf{k}$ . We'll make use of [Theorem 9.8.6](#) to find  $\kappa(t)$ . We have

$$\begin{aligned}\mathbf{r}'(t) &= \langle 1, 2t, e^t \rangle \\ \mathbf{r}''(t) &= \langle 0, 2, e^t \rangle\end{aligned}$$

and so

$$\kappa(t) = \frac{|\langle 2e^t(t-1), -e^t, 2 \rangle|}{|\langle 1, 2t, e^t \rangle|^3} = \frac{\sqrt{4e^{2t}(t-1)^2 + e^{2t} + 4}}{(1+4t^2+e^{2t})^{3/2}}.$$

### 9.8.3 Normal and Binormal Vectors

Consider a curve  $\mathbf{r}(t)$ . Then the direction of the curve is given by the unit tangent vector  $\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|}$ . If we differentiate  $\mathbf{T}(t)$  to get  $\mathbf{T}'(t)$ , we can view this vector as telling us the direction the curve is turning. This leads us to the concept of **unit normal vectors** to a space curve.

**Definition 9.8.8 Unit Normal Vectors.** Consider a space curve given by  $\mathbf{r}(t)$ . The unit normal vector is the vector  $\mathbf{N}(t)$  given by

$$\mathbf{N}(t) = \frac{\mathbf{T}'(t)}{|\mathbf{T}'(t)|},$$

where  $\mathbf{T}'$  is the derivative of the unit tangent vector.

**Example 9.8.9 Unit normal on a circle.** Find the unit normal vector of the curve given by  $\mathbf{r}(t) = \langle \cos t, \sin t \rangle$ .

**Solution.** If we think of a particle moving along  $\mathbf{r}(t)$ , then this particle is just moving along the unit circle. So at every point along this path, the particle should be turning toward the origin in order to stay on the unit circle. So at all points of the curve,  $\mathbf{N}(t)$  should point towards the origin. To prove this, we'll use the formula above to find the unit normal:

$$\begin{aligned}\mathbf{N}(t) &= \frac{\mathbf{T}'(t)}{|\mathbf{T}'(t)|} \\ &= \frac{\langle -\cos t, -\sin t \rangle}{1}\end{aligned}$$

So  $\mathbf{N}(t) = \langle -\cos t, -\sin t \rangle = -\mathbf{r}(t)$ . So at every point of the circle, the unit normal points in the opposite direction of the corresponding position vector, i.e. it points towards the origin.

It's not clear from [Definition 9.8.8](#), but  $\mathbf{N}(t)$  is *always* orthogonal to  $\mathbf{T}(t)$  for any space curve/vector function  $\mathbf{r}(t)$ . We can use these two vectors to get another vector called the **binormal vector**.

**Definition 9.8.10 Binormal Vector.** Let  $\mathbf{r}(t)$  be a vector function with unit tangent  $\mathbf{T}(t)$  and unit normal  $\mathbf{N}(t)$ . Then the unit binormal vector is the vector  $\mathbf{B}(t)$  given by

$$\mathbf{B}(t) = \mathbf{T}(t) \times \mathbf{N}(t).$$

SUGGESTED PROBLEMS: 1, 3, 11-19 odd, 39

## 9.9 Motion in Space

We now move on to examining motion in  $\mathbb{R}^3$ . First, recall that  $\mathbf{r}(t)$  represents the position vector of some particle moving along a curve in space, then its velocity  $\mathbf{v}(t)$  and acceleration  $\mathbf{a}(t)$  are given by

$$\begin{aligned}\mathbf{v}(t) &= \mathbf{r}'(t) \\ \mathbf{a}(t) &= \mathbf{v}'(t) = \mathbf{r}''(t)\end{aligned}$$

The speed of the particle is just the magnitude of the velocity:  $|\mathbf{v}(t)|$ .

**Example 9.9.1 Motion of a projectile.** A projectile is fired out of a cannon with an initial speed of  $200 \frac{\text{m}}{\text{s}}$  to the west and with an angle of elevation of  $30^\circ$ . If the particle was fired from a raised platform that is 50 m off level ground, where does the particle land?

**Solution.** First, we'll assume that  $\mathbf{j}$  points northward and  $\mathbf{k}$  points straight up. Let's assume that the platform is directly above the origin. If we let  $\mathbf{r}(t)$  denote the position (in meters) of the particle at time  $t$  (in seconds), then we can say that  $\mathbf{r}_0 = \mathbf{r}(0) = \langle 0, 0, 50 \rangle$ . We also have

$$\mathbf{v}(0) = -100\sqrt{3}\mathbf{i} + 100\mathbf{k}$$

$$\mathbf{a} = -9.8\mathbf{k}$$

We can integrate up to find the position  $\mathbf{r}(t)$ :

$$\mathbf{v}(t) = -9.8t\mathbf{k} + \mathbf{C}$$

where  $\mathbf{C}$  is an arbitrary constant vector. To find it, we'll use our initial condition on  $\mathbf{v}$ :

$$\mathbf{v}(0) = \mathbf{C} = -100\sqrt{3}\mathbf{i} + 100\mathbf{k}.$$

So  $\mathbf{v}(t) = -9.8t\mathbf{k} - 100\sqrt{3}\mathbf{i} + 100\mathbf{k}$ . Integrating once more to get the position, we have

$$\mathbf{r}(t) = -100\sqrt{3}t\mathbf{i} + (100t - 4.9t^2)\mathbf{k} + \mathbf{D} = -100\sqrt{3}t\mathbf{i} + (100t - 4.9t^2)\mathbf{k} + 50\mathbf{k}.$$

So  $\mathbf{r}(t) = \langle -100\sqrt{3}t, 0, 50 + 100t - 4.9t^2 \rangle$ .

To find where the particle lands, we just set the third component equal to zero and solve for  $t$  to get

$$t = \frac{-100 \pm \sqrt{10000 + 980}}{-9.8} = -.4883, 20.896.$$

We need to choose the positive value for  $t$ , and if we do so we see that when the projectile hits ground it's at position  $\mathbf{r}(20.896) = \langle -3619.38, 0, 0 \rangle$ . So the projectile is a little over 3.5 km to the west.

To describe acceleration on a curve  $\mathbf{r}(t)$ , it can be useful to break it down into **tangential components** and **normal components**.

**Definition 9.9.2 Components of Acceleration.** Let  $\mathbf{r}(t)$  denote a space curve with unit tangent  $\mathbf{T}(t)$ , unit normal  $\mathbf{N}(t)$  and acceleration  $\mathbf{a}$ . The tangential component of acceleration is the scalar  $a_T$  given by

$$a_T = \mathbf{a} \cdot \mathbf{T}.$$

The normal component is the scalar  $a_N$  given by

$$a_N = \mathbf{a} \cdot \mathbf{N}.$$

$a_T$  represents how much of the acceleration is directed tangent to the curve  $\mathbf{r}(t)$ , whereas  $a_N$  represents how much of the acceleration is directed perpendicular to the tangent.

**Example 9.9.3 Components of acceleration for a projectile.** Consider the particle from [Example 9.9.1](#). At what point is the tangential component of acceleration greatest?

**Solution.** Since  $\mathbf{a} = -9.8\mathbf{k}$ , the tangential component should be greatest when the projectile is fired or when it hits the ground, since these are the points where the direction of the trajectory most closely matches the direction of acceleration. To actually verify this, we'll find the tangential component using [Definition 9.9.2](#). Since

$$\mathbf{T}(t) = \left\langle -\frac{100\sqrt{3}}{\sqrt{(100 - 9.8t)^2 + 30000}}, 0, \frac{100 - 9.8t}{\sqrt{(100 - 9.8t)^2 + 30000}} \right\rangle$$

it follows that

$$a_T = \mathbf{a} \cdot \mathbf{T}$$

$$= -9.8 \frac{100 - 9.8t}{\sqrt{(100 - 9.8t)^2 + 30000}}$$

So

$$\begin{aligned} |a_T| &= 9.8 \frac{|100 - 9.8t|}{\sqrt{(100 - 9.8t)^2 + 30000}} \\ &= \frac{9.8}{\sqrt{1 + \frac{30000}{(100 - 9.8t)^2}}} \end{aligned}$$

We can make this as large as possible by making  $(100 - 9.8t)^2$  as large as possible, and this reaches its largest value at  $t = 20.896$ , when the projectile hits the ground.

```
t = var('t')
r = vector([-100*sqrt(3)*t, 0, 50 + 100*t - 4.9*t^2])
T = derivative(r,t).normalized()
show(T)
```

SUGGESTED PROBLEMS: 1-11 odd, 19, 31



# Chapter 10

## Partial Derivatives

For functions that depend on a single independent variable such as  $x$  or  $t$ , derivatives are useful because they represent how the function itself changes as the independent variable changes. However, there are many situations where it makes sense to consider functions of more than one independent variable. Therefore, we will want to extend the notion of derivative to include these functions as well. Our primary tool for this will be the **partial derivative**, which we'll see in [Section 10.3](#). But we'll need to build up some terminology to get there.

### 10.1 Functions of Several Independent Variables

#### 10.1.1 Functions of Two Variables

A function of several variables is a function that depends on more than one independent variable. One such example comes from differential equations, where quantities such as the temperature of a metal rod depend on both a position  $x$  and a time  $t$ .

**Definition 10.1.1 Functions of Two Variables.** A **function of two variables** is a rule  $f$  or  $z = f(x, y)$  that assigns pairs  $(x, y)$  of real numbers to real numbers. The **domain** is the set  $D$  in  $\mathbb{R}^2$  for which the rule is defined. The **range** is the set  $R$  in  $\mathbb{R}$  of all possible values of  $f$ .  $x, y$  are called independent variables and  $z$  is called the **dependent variable**.

**Example 10.1.2 Domain of a function.** Let  $f(x, y) = \frac{1}{\sqrt{4-x^2-y^2}}$ . Find the domain of  $f$ .

**Solution.** We'll break this down into pieces. First, since we're dealing with square roots in real numbers we need  $x^2 + y^2 \leq 4$ . Second, we need to make sure that  $\sqrt{4-x^2-y^2} \neq 0$ . So the domain of  $f$  is the set of all points  $(x, y)$  in  $\mathbb{R}^2$  such that  $x^2 + y^2 < 4$ . In other words, the domain is just the inside of the circle of radius 2 centered at the origin.

We can also graph functions of two variables by viewing  $f(x, y)$  as representing the height of some surface over the point  $(x, y)$  in the  $xy$ -plane. In other words, just as  $y = f(x)$  represents a curve in  $\mathbb{R}^2$ , the equation  $z = f(x, y)$  represents a surface in  $\mathbb{R}^3$ .

**Example 10.1.3 Sketching a cone.** Sketch the surface given by  $f(x, y) = \sqrt{4x^2 + y^2}$ .

**Solution.** We need to sketch  $z = \sqrt{4x^2 + y^2}$ . We can start by squaring both sides to get  $z^2 = 4x^2 + y^2$ . If we look at the right hand side, it looks a lot like an ellipse, so we'll try to work with that. Assuming  $z \neq 0$ , we can rewrite the equation to get

$$\frac{x^2}{z^2/4} + \frac{y^2}{z^2} = 1$$

For any given positive value of  $z$ , this represents an ellipse that extends to  $\frac{z}{2}$  in the  $x$  direction and  $z$  in the  $y$  direction. So as  $z$  increases, these ellipses increase in size as well.

### 10.1.2 Contour Plots

**Definition 10.1.4 Contour Plots.** Given a function  $f(x, y)$ , a **contour plot** is the plot of an equation of the form  $f(x, y) = k$  for some constant  $k$ .

A contour plot can be viewed as slicing a surface in  $\mathbb{R}^3$  with a level plane. The contour plot is exactly where the plane intersects the surface.

**Example 10.1.5 Contours on a paraboloid.** What do the contours of  $f(x, y) = x^2 + \frac{y^2}{9}$  look like?

**Solution.** If we pick different (nonnegative) values for  $k$ , then all of the contours of  $f(x, y)$  take the form

$$x^2 + \frac{y^2}{9} = k$$

which turns into

$$\frac{x^2}{k} + \frac{y^2}{9k} = 1.$$

This is a family of ellipses.

SUGGESTED PROBLEMS: 1, 5, 11, 13

## 10.2 Limits and Continuity

### 10.2.1 Limits

Just as functions of one variable have a notion of limit, the same can be said for functions of several variables. However, limits involving several variables tend to be more restrictive.

**Definition 10.2.1 Limit of a Function of Two Variables.** Let  $f(x, y)$  be some function and let  $(a, b)$  be a point arbitrarily close to the domain of  $f$ . We say that

$$\lim_{(x,y) \rightarrow (a,b)} f(x, y) = L$$

if  $f(x, y)$  gets arbitrarily close to  $L$  as  $(x, y)$  gets arbitrarily close to  $(a, b)$ .

Here's why limits can be tricky in two dimensions: for the limit to exist,  $f(x, y)$  must approach the same value  $L$  no matter how  $(x, y)$  approaches  $(a, b)$ !

**Example 10.2.2.** Find

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 \sin y}{x^3 + 2y^3}$$

or show that it doesn't exist.

**Solution.** First, we'll start simple. If we set  $x = 0$  then the limit becomes

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 \sin y}{x^3 + 2y^3} = \lim_{y \rightarrow 0} \frac{0}{2y} = 0.$$

Likewise, if we set  $y = 0$  then the limit becomes

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 \sin y}{x^3 + 2y^3} = \lim_{x \rightarrow 0} \frac{0}{x^3} = 0.$$

From this, we might suspect that the limit actually exists. However, we need to make sure that this approaches 0 along *all possible paths* to  $(0, 0)$ . Here's another path we can try:  $y = x$ . If we do this, then the limit becomes

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 \sin y}{x^3 + 2y^3} = \lim_{x \rightarrow 0} \frac{x^2 \sin x}{3x^3} = \frac{1}{3}.$$

Since the value that  $\frac{x^2 \sin y}{x^3 + 2y^3}$  approaches as  $(x, y) \rightarrow (0, 0)$  depends on how  $(x, y)$  approaches  $(0, 0)$ , we can say that the limit does not exist.

In general, showing limits exist for functions of several variables can be tricky. However, occasionally we can use some tricks to make these problems more manageable.

**Example 10.2.3 Showing a limit exists.** Find  $\lim_{(x,y) \rightarrow (0,0)} (x^2 + y^2) \ln(x^2 + y^2)$ .

**Solution.** Since the inside of the limit depends on  $x^2 + y^2$ , this suggests that switching to polar coordinates might be beneficial. So let's try that!

$$\lim_{(x,y) \rightarrow (0,0)} (x^2 + y^2) \ln(x^2 + y^2) = \lim_{(r,\theta) \rightarrow (0,0)} r^2 \ln r^2.$$

Here's why this helps us. If we look at our new limit, we see that it doesn't depend on  $\theta$  at all. In other words, *it doesn't depend on the direction* we approach the origin from, just the distance from the origin  $r$ . So

$$\begin{aligned} \lim_{(r,\theta) \rightarrow (0,0)} r^2 \ln r^2 &= \lim_{r \rightarrow 0^+} 2r^3 \ln r \\ &= 0 \end{aligned}$$

and so the original limit also equals 0.

### 10.2.2 Continuous Functions

Now that we have the notion of limit, we can define continuous functions of two variables.

**Definition 10.2.4 Continuous Function of Two Variables.** Let  $f(x, y)$  be a function of two variables. We say that  $f$  is **continuous** at  $(a, b)$  if

$$\lim_{(x,y) \rightarrow (a,b)} f(x, y) = f(a, b).$$

If  $f$  is continuous at all points in a set  $D$ , we say that  $f$  is continuous on  $D$ .

We won't typically use the definition of continuity in this course. We'll just make of several common types of continuous functions:

1. Polynomials in two variables, such as  $x - 3x^4y^5$ , are continuous on  $\mathbb{R}^2$ .

2. Ratios of polynomials (i.e. **rational functions**), such as  $\frac{y^3-x}{1+x^{20}y^5}$ , are continuous on their domains.
3. Compositions of the form  $g(f(x, y))$ , where  $g(x)$  and  $f(x, y)$  are continuous, are continuous wherever they are defined.

**Example 10.2.5 Limit of a continuous function.** Let  $f(x, y) = \tan \frac{3x^2-5y}{1-xy}$ . Find

$$\lim_{(x,y) \rightarrow (3,4)} f(x, y).$$

**Solution.** First, note that this is a composition of continuous functions, so  $f(x, y)$  should be continuous wherever it's defined. Since this is defined at  $(3, 4)$ , we can say that  $f$  is continuous there as well and that

$$\lim_{(x,y) \rightarrow (3,4)} f(x, y) = f(3, 4) = \frac{7}{-11}.$$

SUGGESTED PROBLEMS: 3, 11, 19, 23, 31

## 10.3 Partial Derivatives

Now that we know how to take limits for functions of two variables, we can define a notion of derivative for these functions as well, which we'll call the **partial derivative**. The main idea behind the partial derivative is to measure how a function  $f(x, y)$  changes as a single variable changes. Although the definition given is only for functions of two variables, it also extends to functions of many variables.

**Definition 10.3.1 Partial Derivatives.** Let  $f(x, y)$  be a function of two variables. The **partial derivative of  $f$  with respect to  $x$**  is the function

$$\frac{\partial f}{\partial x} = f_x(x, y) = \lim_{h \rightarrow 0} \frac{f(x+h, y) - f(x, y)}{h},$$

assuming that this limit exists. Similarly, the **partial derivative of  $f$  with respect to  $y$**  is the function

$$\frac{\partial f}{\partial y} = f_y(x, y) = \lim_{h \rightarrow 0} \frac{f(x, y+h) - f(x, y)}{h},$$

assuming that this limit exists.

When computing partial derivatives with respect to  $x$ , you just need to differentiate  $f(x, y)$  treating it as a function of  $x$  alone and considering  $y$  as a constant. Similarly, if you want to compute  $f_y$  you differentiate with respect to  $y$  and treat  $x$  as constant.

**Example 10.3.2.** Let  $f(x, y) = ye^{x-y} - x^3 \tan y^4$ . Find  $\frac{\partial f}{\partial y}$ .

**Solution.** Since we need to find  $\frac{\partial f}{\partial y}$ , we differentiate  $f(x, y)$  with respect to  $y$  and treat  $x$  as a constant:

$$\begin{aligned} \frac{\partial f}{\partial y} &= \frac{\partial}{\partial y} [ye^{x-y} - x^3 \tan y^4] \\ &= e^{x-y} - ye^{x-y} - 4x^3 y^3 \sec^2 y^4 \end{aligned}$$

Geometrically,  $f_x(a, b)$  represents the rate of change of the surface  $z = f(x, y)$  at the point  $(a, b)$  in the  $x$  direction, and a similar statement is true for  $f_y(a, b)$ .

**Example 10.3.3 Partial derivatives on the unit sphere.** Let  $z = \sqrt{1 - x^2 - y^2}$ .

Where is  $\frac{\partial z}{\partial y}$  equal to 0?

**Solution.** We can find this algebraically, but we'll try to answer this geometrically instead. First, note that  $z = \sqrt{1 - x^2 - y^2}$  is actually the top half of the unit sphere. If we're trying to find where  $\frac{\partial z}{\partial y}$  is zero, then we need to find where this surface is "flat" when moving in the  $y$  direction. If we think about this a bit, this should only occur when  $y = 0$ , i.e. along the  $x$ -axis. At any other location on the unit sphere, moving in the  $y$ -direction on the unit sphere requires going uphill or downhill, which means  $z_y \neq 0$  at these locations.

We can also find higher partial derivatives. For example, the second partial derivatives of a function  $f(x, y)$  are given by

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial x} \right) = (f_x)_x \\ \frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial y} \right) = (f_y)_y \\ \frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) = (f_y)_x \\ \frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right) = (f_x)_y\end{aligned}$$

The last two derivatives are examples of **mixed derivatives**.

**Example 10.3.4 Mixed partials.** Let  $f(x, y) = \sin^2(mx + ny)$  where  $m, n$  are constants. Find the mixed partials  $f_{xy}$  and  $f_{yx}$ .

**Solution.** We have

$$\begin{aligned}f_{xy} &= (f_x)_y \\ &= (2m \cos(mx + ny) \sin(mx + ny))_y \\ &= -2mn \sin^2(mx + ny) + 2mn \cos^2(mx + ny) \\ &= 2mn \cos(2mx + 2ny)\end{aligned}$$

as well as

$$\begin{aligned}f_{yx} &= (f_y)_x \\ &= (2n \cos(mx + ny) \sin(mx + ny))_x \\ &= -2mn \sin^2(mx + ny) + 2mn \cos^2(mx + ny) \\ &= 2mn \cos(2mx + 2ny)\end{aligned}$$

In the last example the mixed partial were equal, but this is not always true. However, if  $f(x, y)$  is a function with "nice" mixed partial derivatives, then its mixed partials will be equal.

**Theorem 10.3.5 Clairaut's Theorem.** Suppose that  $f(x, y)$  is defined on a disk  $D$  that contains the point  $(a, b)$ . If  $f_{xy}$  and  $f_{yx}$  are both continuous on  $D$ , then

$$f_{xy}(a, b) = f_{yx}(a, b).$$

In other words, the mixed partials are equal to each other wherever they happen to be continuous.

Partial derivatives are important for their use in **partial differential equations (PDEs)**, which are equations that involve partial derivatives of an unknown function  $f$ .

**Example 10.3.6 The Wave Equation.** One of the most important PDEs is the **wave equation**

$$u_{tt} = c^2 u_{xx},$$

which is used to model the displacement of a string  $u(x, t)$  at position  $x$  and time  $t$ . Determine if

$$u(x, t) = \sin(x - 3t) + \ln(x + 3t)$$

is a solution of the wave equation.

**Solution.** We need to plug  $u$  into the PDE and check that it satisfies the PDE. Since

$$\begin{aligned} u_{tt} &= -9 \cos(x - 3t) - \frac{9}{(x + 3t)^2} \\ u_{xx} &= -\sin(x - 3t) - \frac{1}{(x + 3t)^2} \end{aligned}$$

it follows that  $u_{tt} = 9u_{xx}$ . So  $u(x, t)$  satisfies the wave equation with  $c = 3$ , and so is a solution of the wave equation.

SUGGESTED PROBLEMS: 1, 3, 7, 11, 19, 31, 45, 49, 51, 53, 61

## 10.4 Tangent Planes and Linear Approximations

### 10.4.1 Tangent Planes

Recall that if a curve  $y = f(x)$  is differentiable at a point  $x_0$ , then it has a tangent line passing through  $(x_0, f(x_0))$ . The tangent line can be viewed as a linear approximation of the curve  $y = f(x)$  near the point  $x_0$ . We can apply a similar ideas to surfaces  $z = f(x, y)$ . It turns out that if we have a surface  $z = f(x, y)$  and a point  $(x_0, y_0, f(x_0, y_0))$  on the surface, then every tangent vector at this point is contained in a single plane called the **tangent plane**.

We'd like to find the equation of this plane. First, recall that every plane can be described by an equation of the form

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0.$$

We can view  $x_0, y_0$  and  $z_0 = f(x_0, y_0)$  as given, so we just need to find  $a, b, c$ . If we assume that  $c \neq 0$ , then we can rewrite this equation to obtain

$$z - z_0 = A(x - x_0) + B(y - y_0)$$

where  $A = -\frac{a}{c}$  and  $B = -\frac{b}{c}$ . If we set  $y = y_0$ , then we have  $z - z_0 = A(x - x_0)$ . This is the equation of a line tangent to the surface and parallel to the  $x$ -axis, and so the slope of this line must be

$$A = f_x(x_0, y_0)$$

since the slope of a tangent line in the  $x$ -direction gives the rate of change in the  $x$ -direction. Similarly,

$$B = f_y(x_0, y_0).$$

Putting all of this together gives the following theorem.

**Theorem 10.4.1 Tangent Planes to Surfaces.** Let  $z = f(x, y)$  be a surface and suppose that  $f$  has continuous partial derivatives at the point  $(x_0, y_0)$ . Then the tangent plane to the surface  $z = f(x, y)$  at the point  $(x_0, y_0)$  is given by

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0),$$

where  $z_0 = f(x_0, y_0)$ .

**Example 10.4.2 Approximations by tangent planes.** Find the tangent plane to  $f(x, y) = xe^{xy}$  at the point  $(2, 0, 2)$ . Use this to approximate  $f(1.9, 0)$

**Solution.** The equation of the tangent plane is given by

$$z - 2 = (x - 2) + 4(y - 0)$$

which we can rewrite as

$$z = x + 4y.$$

So at  $(1.9, 0)$ , we should have

$$f(1.9, 0) \approx 1.9 + 4(0) = 1.9.$$

## 10.4.2 Linear Approximations

[Example 10.4.2](#) shows that we can use tangent planes to approximate complicated functions. This leads us to the idea of a linear approximation, or **linearization** of a function of the form  $f(x, y)$ .

**Definition 10.4.3 Linearization.** Let  $f(x, y)$  be a function for which  $f_x$  and  $f_y$  are both continuous at a point  $(x_0, y_0)$ . Then the linearization of  $f(x, y)$  at  $(x_0, y_0)$  is the function  $L(x, y)$  given by

$$L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

For well-behaved functions (i.e. functions that have continuous partial derivatives),  $L(x, y) \approx f(x, y)$  if  $(x, y)$  is close to the point  $(x_0, y_0)$ .

**Example 10.4.4 Linearization of an exponential and a sinusoid.** Let  $u(x, t) = 30e^{-t} \sin x$ . Find the linearization of  $u$  at the point  $(0, \frac{\pi}{2})$ .

**Solution.** By [Definition 10.4.3](#), the linearization is given by

$$L(x, y) = 30 - 30(t - 0) = 30(1 - t).$$

[Definition 10.4.3](#) can also be extended to functions with more than two variables.

**Example 10.4.5 Linearization in three variables.** Let  $f(x, y, z) = xze^{-y^2-z^2}$ . Find the linearization  $L(x, y, z)$  at the point  $(-3, \sqrt{\ln 2}, -\sqrt{\ln 3})$ .

**Solution.** The formula we need to use now is

$$L(x, y, z) = f_x(-3, \sqrt{\ln 2}, -\sqrt{\ln 3})(x+3) + f_y(-3, \sqrt{\ln 2}, -\sqrt{\ln 3})(y-\sqrt{\ln 2}) + f_z(-3, \sqrt{\ln 2}, -\sqrt{\ln 3})(z+\sqrt{\ln 3}).$$

SUGGESTED PROBLEMS: 1, 15, 19

## 10.5 The Chain Rule in Several Variables

Recall that from Calc I, we know that

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

This is just the chain rule, which tells us how to differentiate functions that depend on other functions. Or to think of it another way, it tells us how to differentiate *variables* like  $f$  that depend on other variables like  $g$  and  $x$ . What we'd like to do now is to extend the chain rule to more complicated situations. We'll start with the following formula.

**Theorem 10.5.1 Chain Rule for One Independent Variable.** *Let  $z = f(x, y)$  be differentiable (so the partial derivatives  $f_x$  and  $f_y$  exist). If  $x$  and  $y$  both depend on  $t$  (i.e.  $x = x(t)$  and  $y = y(t)$ ), then*

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt}$$

See [Figure 10.5.2](#).

**Figure 10.5.2:** The chain rule for one independent variable.

**Example 10.5.3 Using the chain rule.** Let  $f(x, y, z) = \ln \sqrt{x^2 + y^2 + z^2}$  where

$$x = \sin t, \quad y = \cos t \quad \text{and} \quad z = \tan t.$$

Find  $\left. \frac{df}{dt} \right|_{t=0}$ .

**Solution.** First, let's try to simplify  $f$ . If we do so, we get

$$f(x, y, z) = \frac{1}{2} \ln(x^2 + y^2 + z^2).$$

Now, to find  $\frac{df}{dt}$  we need to use the chain rule, which gives

$$\begin{aligned} \frac{df}{dt} &= f_x x'(t) + f_y y'(t) + f_z z'(t) \\ &= \frac{x}{x^2 + y^2 + z^2} \cos t - \frac{y}{x^2 + y^2 + z^2} \sin t + \frac{z}{x^2 + y^2 + z^2} \sec^2 t \\ &= \frac{\sin t \cos t}{\sec^2 t} - \frac{\cos t \sin t}{\sec^2 t} + \frac{\tan t}{\sec^2 t} \\ &= \cos t \sin t \\ &= \frac{1}{2} \sin 2t \end{aligned}$$

So

$$\left. \frac{df}{dt} \right|_{t=0} = 0.$$

If we have more than one independent variable, then the situation becomes a little bit more complicated.



**Theorem 10.5.4 Chain Rule with Two Independent Variables.** Let  $f(x, y)$  be differentiable, and suppose that  $x = x(s, t)$  and  $y = y(s, t)$ . Then

$$\begin{aligned}\frac{\partial f}{\partial s} &= f_x x_s + f_y y_s \\ \frac{\partial f}{\partial t} &= f_x x_t + f_y y_t\end{aligned}$$



**Figure 10.5.5:** The chain rule for two independent variables.

**Example 10.5.6 Chain rule with two independent variables.** Let  $f(x, y, z) = x^2 - \sin^{-1}(z - y)$ , where

$$x = e^t, \quad y = s + t \quad \text{and} \quad z = t^2 - st.$$

Find  $\frac{\partial f}{\partial s}$ .

**Solution.** A good way to tackle these problems is to set up a tree diagram like [Figure 10.5.5](#) to list the dependencies between all of the variables. For  $f(x, y, z)$ , we get



So we can say that

$$\begin{aligned}\frac{\partial f}{\partial s} &= f_y y_s + f_z z_s \\ &= \frac{1}{\sqrt{1 - (z - y)^2}} - \frac{1}{\sqrt{1 - (z - y)^2}}(-t) \\ &= \frac{1 + t}{\sqrt{1 - (z - y)^2}} \\ &= \frac{1 + t}{\sqrt{1 - (t^2 - st - s - t)^2}}\end{aligned}$$

**Example 10.5.7 Chain rule from a tree diagram.** Let  $R = f(x, y, z, t)$ , where  $x = x(u, v, w)$ ,  $y = y(u, v)$ ,  $z = z(v)$  and  $t = t(v, w)$ . Assume that all of these functions are differentiable. Find a formula for  $\frac{\partial R}{\partial w}$ .

**Solution.** We can set up a tree diagram as above to find the right formula for  $R_w$ . If we do so, we obtain

$$R_w = R_x x_w + R_t t_w.$$

SUGGESTED PROBLEMS: 1, 7, 11, 15, 21

## 10.6 Directional Derivatives and Gradients

For a differentiable function  $f(x, y)$ , the partial derivatives  $f_x$  and  $f_y$  give the rates of change of  $f$  in the directions of the  $x$ -axis and  $y$ -axis, respectively. To put this another way,  $f_x$  gives the rate of change of  $f$  in the direction of the vector  $\mathbf{i}$ , and  $f_y$  gives the rate of change in the direction of the vector  $\mathbf{j}$ . But we have infinitely many directions to work with. How can we find the rate of change of  $f$  in *any* direction? We'll look at two tools to help us with this: directional derivatives and gradients.

### 10.6.1 The Directional Derivative

Suppose we have a function  $f(x, y)$  defined at a point  $(x_0, y_0)$ . We want to find how quickly  $f(x, y)$  is changing in the direction of some unit vector  $\mathbf{u} = \langle a, b \rangle$ . We can estimate this by moving from the point  $(x_0, y_0)$  by some small increment  $h\mathbf{u}$ , and evaluating  $f$  at this new point, which is just  $(x_0 + ha, y_0 + hb)$ . So the approximate rate of change in  $f$  in the direction of  $\mathbf{u}$  is given by

$$\frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

As  $h \rightarrow 0$ , this should become exact. This leads us to our next definition.

**Definition 10.6.1 Directional Derivative.** The **directional derivative** of the function  $f(x, y)$  at the point  $(x_0, y_0)$  in the direction of the vector  $\mathbf{u} = \langle a, b \rangle$  is given by

$$D_{\mathbf{u}}f(x_0, y_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

assuming this limit exists. As we mentioned above,  $D_{\mathbf{i}}f = f_x$  and  $D_{\mathbf{j}}f = f_y$ .

We have a definition for the directional derivative, but we don't want to have to use this all the time. So we'd like to get a more practical formula.

**Theorem 10.6.2 Computing the Directional Derivative.** Let  $f(x, y)$  be a function whose partial derivatives  $f_x$  and  $f_y$  exist and are continuous. Let  $\mathbf{u} = \langle a, b \rangle$  be a unit vector. Then  $D_{\mathbf{u}}f(x, y)$  exists, and is given by

$$D_{\mathbf{u}}f(x, y) = f_x(x, y)a + f_y(x, y)b.$$

*Proof.* First, set

$$g(h) = f(x_0 + ha, y_0 + hb)$$

Then by the Calculus I definition of the derivative,

$$g'(0) = \lim_{h \rightarrow 0} \frac{g(h) - g(0)}{h} = \lim_{h \rightarrow 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h} = D_{\mathbf{u}}f(x_0, y_0).$$

On the other hand, if we set  $x = x_0 + ha$  and  $y = y_0 + hb$ , then  $g(h) = f(x, y)$ . We can then find  $g'(h)$  using the chain rule from [Section 2.5](#) to get

$$g'(h) = f_x \frac{dx}{dh} + f_y \frac{dy}{dh} = f_x(x, y)a + f_y(x, y)b$$

So  $g'(0) = f_x(x_0, y_0)a + f_y(x_0, y_0)b$ . Therefore

$$D_{\mathbf{u}}f(x_0, y_0) = f_x(x_0, y_0)a + f_y(x_0, y_0)b. \quad \square$$

The formula in [Theorem 10.6.2](#) extends to functions such as  $f(x, y, z)$  as well. We just need to include more partial derivatives.

**Example 10.6.3 Computing a directional derivative.** Let  $f(r, s) = \frac{r}{r^2 + s^2}$ . Find the directional derivative of  $f$  in the direction of  $\mathbf{v} = 3\mathbf{i} + 5\mathbf{j}$  at the point  $(1, 2)$ .

**Solution.** We can use [Theorem 10.6.2](#) to find the directional derivative. However, there's something we need to watch out for here. Both [Definition 10.6.1](#) and [Theorem 10.6.2](#) require using a *unit vector*, but our vector  $\mathbf{v}$  isn't a unit vector at all. So we need to **normalize**  $\mathbf{v}$  to get a unit vector  $\mathbf{u}$  that is parallel to  $\mathbf{v}$ :

$$\mathbf{u} = \frac{\mathbf{v}}{|\mathbf{v}|} = \left\langle \frac{3}{\sqrt{34}}, \frac{5}{\sqrt{34}} \right\rangle$$

Since

$$f_r = \frac{s^2 - r^2}{(r^2 + s^2)^2}$$

$$f_s = -\frac{2rs}{(r^2 + s^2)^2}$$

we see that the directional derivative we need is given by

$$\begin{aligned} D_{\mathbf{u}}f(1, 2) &= f_r(1, 2) \frac{3}{\sqrt{34}} + f_s(1, 2) \frac{5}{\sqrt{34}} \\ &= \frac{3}{25} \frac{3}{\sqrt{34}} - \frac{4}{25} \frac{5}{\sqrt{34}} \end{aligned}$$

## 10.6.2 Gradients

[Theorem 10.6.2](#) shows that the partial derivatives of a function are important in computing the directional derivatives. This leads us to the concept of a **gradient**.

**Definition 10.6.4 Gradient of a Function.** Let  $f(x, y, z)$  be a function with partial derivatives  $f_x, f_y, f_z$ . The gradient of  $f$  is the vector  $\nabla f = \langle f_x, f_y, f_z \rangle$ .

**Example 10.6.5 Computing a gradient.** Let  $f(x, y) = x^2 + y^2$ . Compute  $\nabla f$ .

**Solution.** We have

$$\nabla f = \langle 2x, 2y \rangle$$

We can restate [Theorem 10.6.2](#) very quickly in terms of the gradient. In particular, if  $\mathbf{u}$  is a vector (not necessarily a unit vector) then

$$D_{\mathbf{u}}f = \nabla f \cdot \frac{\mathbf{u}}{|\mathbf{u}|} \quad (10.6.1)$$

**Example 10.6.6 Directional derivative using the gradient.** Is the function  $f(x, y, z) = xy + yz + zx$  increasing or decreasing at  $P(1, -1, 3)$  and in the direction of  $Q(2, 4, 5)$ .

**Solution.** We need to compute  $D_{\vec{PQ}}f(1, -1, 3)$  and check if it's positive or negative. Since  $\nabla f = \langle y, z, x \rangle$  and  $\vec{PQ} = \langle 1, 5, 2 \rangle$ , it follows that

$$\begin{aligned} D_{\vec{PQ}}f(1, -1, 3) &= \nabla f(1, -1, 3) \cdot \frac{\vec{PQ}}{|\vec{PQ}|} \\ &= \langle -1, 3, 1 \rangle \cdot \left\langle \frac{1}{\sqrt{30}}, \frac{5}{\sqrt{30}}, \frac{2}{\sqrt{30}} \right\rangle \\ &= \frac{16}{\sqrt{30}} \end{aligned}$$

Hence  $f$  is increasing at  $(1, -1, 3)$  and in the direction of  $(2, 4, 5)$ .

The gradient is also useful for telling us where a function is increasing and decreasing the greatest. By [\(10.6.1\)](#), we can write

$$D_{\mathbf{u}}f = \nabla f \cdot \frac{\mathbf{u}}{|\mathbf{u}|} = |\nabla f| \cos \theta,$$

where  $\theta$  is the acute angle between  $\nabla f$  and  $\mathbf{u}$ . It follows that

$$-|\nabla f| \leq D_{\mathbf{u}}f \leq |\nabla f|.$$

We summarize this in the following theorem.

**Theorem 10.6.7 Direction of Greatest Increase and Decrease.** Let  $f$  be a function with continuous first partial derivatives. Then  $\nabla f$  points in the direction of greatest increase of  $f$ , and the rate of maximum increase of  $f$  is given by  $|\nabla f|$ . Similarly,  $-\nabla f$  points in the direction of greatest decrease, and the rate of maximum decrease of  $f$  is given by  $|\nabla f|$ .

**Example 10.6.8 Finding the direction and rate of maximum decrease.**

Let  $f(s, t) = te^{st}$ . What is the direction and rate of greatest decrease of  $f$  at the point  $(0, 2)$ ?

**Solution.** We need to compute  $-\nabla f$ , which gives

$$-\nabla f = \langle -(1 + st)e^{st}, -ste^{st} \rangle$$

So the direction of greatest decrease at  $(0, 2)$  is given by  $-\nabla f(0, 2) = \langle -1, 0 \rangle$ . The rate of maximum decrease in this direction is  $-|\nabla f(0, 2)| = -|\langle -1, 0 \rangle| = -1$ .

SUGGESTED PROBLEMS: 1, 3, 11, 15, 21

## 10.7 Extreme Values

### 10.7.1 Local Maxima and Minima

One of the most important applications of the derivative in Calculus I is in finding local maxima and minima of functions  $y = f(x)$ . We'd like to extend this concept to functions of multiple variables, but first we need to define what local maxima and minima actually are in this new setting.

**Definition 10.7.1 Local Maxima and Minima.** We say that a function  $f(x, y)$  has a local maximum (respectively, local minimum) at  $(x_0, y_0)$  if  $f(x, y) \leq f(x_0, y_0)$  (respectively,  $f(x, y) \geq f(x_0, y_0)$ ) for all points  $(x, y)$  contained in some disk centered at  $(x_0, y_0)$ .

Just as we used derivatives in Calculus I to find maxima and minima, we can use partial derivatives for that purpose here.

**Theorem 10.7.2 Derivatives at a Local Maximum or Minimum.** Suppose that  $f(x, y)$  is differentiable, and has a local maximum or minimum at  $(x_0, y_0)$ . Then

$$f_x(x_0, y_0) = f_y(x_0, y_0) = 0.$$

**Theorem 10.7.2** does not tell us if a differentiable function  $f(x, y)$  has a local maximum or minimum at  $(x_0, y_0)$ , but it does tell us where to look. We call  $(x_0, y_0)$  a **critical point** of  $f(x, y)$  if  $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$  or if these partial derivatives fail to exist at  $(x_0, y_0)$ .

**Example 10.7.3 Critical points of an exponential.** Let  $f(x, y) = xe^{-x^2-y^2}$ . Find its critical points.

**Solution.** Find the critical points by solving  $f_x = 0$  and  $f_y = 0$ . Since

$$f_x = e^{-x^2-y^2} - 2x^2e^{-x^2-y^2} = e^{-x^2-y^2}(1 - 2x^2)$$

then  $f_x = 0$  forces  $x = \pm\sqrt{\frac{1}{2}}$ . Similarly, since

$$f_y = -2xye^{-x^2-y^2}$$

then  $f_y = 0$  forces  $x = 0$  or  $y = 0$ . So the only points that satisfy both  $f_x = 0$  and  $f_y = 0$  are  $(\pm\sqrt{\frac{1}{2}}, 0)$ . So these must be our critical points.

We found the critical points in [Example 10.7.3](#) without too much trouble, but these are still not guaranteed to be local maxima or minima. To determine if they are, we need to use the **second derivative test**.

**Theorem 10.7.4 Second Derivative Test.** *Let  $f(x, y)$  be a differentiable function with continuous second partial derivatives, and let  $(x_0, y_0)$  be a critical point of  $f$ . Define  $D$  to be*

$$D(x, y) = f_{xx}f_{yy} - (f_{xy})^2$$

*Then we have the following possibilities:*

- If  $D(x_0, y_0) > 0$  and  $f_{xx}(x_0, y_0) > 0$ , then  $(x_0, y_0)$  is a local minimum.
- If  $D(x_0, y_0) > 0$  and  $f_{xx}(x_0, y_0) < 0$ , then  $(x_0, y_0)$  is a local maximum.
- If  $D(x_0, y_0) < 0$ , then  $(x_0, y_0)$  is neither a local minimum nor local maximum. In this case we call  $(x_0, y_0)$  a **saddle point**.
- If  $D(x_0, y_0) = 0$ , the test is inconclusive.

The quantity  $D$  from [Theorem 10.7.4](#) is actually a determinant:

$$D(x, y) = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix}$$

**Example 10.7.5 Local extrema of an exponential.** Classify the critical points of the function from [Example 10.7.3](#).

**Solution.** We have  $f(x, y) = xe^{-x^2-y^2}$ , and the critical points are  $(-\sqrt{\frac{1}{2}}, 0)$  and  $(\sqrt{\frac{1}{2}}, 0)$ . We need to find  $D(x, y)$ . Since  $f_x = (1 - 2x^2)e^{-x^2-y^2}$ , we get

$$\begin{aligned} f_{xx} &= -4xe^{-x^2-y^2} - 2x(1 - 2x^2)e^{-x^2-y^2} \\ f_{xy} &= -2y(1 - 2x^2)e^{-x^2-y^2} = f_{yx} \end{aligned}$$

and similarly

$$f_{yy} = -2xe^{-x^2-y^2} + 4xy^2e^{-x^2-y^2}$$

At  $(-\sqrt{\frac{1}{2}}, 0)$ , we have

$$\begin{aligned} D\left(-\sqrt{\frac{1}{2}}, 0\right) &= 4\sqrt{\frac{1}{2}}e^{-\frac{1}{2}} \cdot 2\sqrt{\frac{1}{2}}e^{-\frac{1}{2}} \\ &> 0 \end{aligned}$$

and  $f_{xx}(-\sqrt{\frac{1}{2}}, 0) > 0$  as well. So this is a local minimum. At  $(\sqrt{\frac{1}{2}}, 0)$ , we still have  $D > 0$ , but now  $f_{xx} < 0$ . So this is a local maximum.

## 10.7.2 Absolute Maxima and Minima

Even if a function  $f(x, y)$  has a local minimum at  $(x_0, y_0)$ , there's no guarantee that this is the smallest value the function can take on its domain. A similar statement is also true for local maxima, which leads to the following definition.

**Definition 10.7.6 Absolute Maxima and Minima.** Let  $f(x, y)$  be a function defined on the set  $D$ . We call the point  $(x_0, y_0)$  an **absolute minimum** (respectively, **absolute maximum**) if  $f(x, y) \geq f(x_0, y_0)$  (respectively,  $f(x, y) \leq f(x_0, y_0)$ ) for all points  $(x, y)$  in  $R$ .

As a first example, consider  $f(x, y) = \sqrt{1 - x^2 - y^2}$ , which gives the top half of the unit sphere. Using critical points and the second derivative test, we can show that the origin is a local maximum. However, this method is blind to the absolute minimum values of  $f(x, y)$ , which occur at the boundary  $x^2 + y^2 = 1$ , where the sphere intersects the  $xy$ -plane. So this means we can't use critical points alone to find absolute maxima and minima, in general. However, the following theorem says that they'll get us most of the way there.

**Theorem 10.7.7 Extreme Value Theorem.** *Let  $f(x, y)$  be a continuous function defined on a closed and bounded set  $D$ .<sup>1</sup> Then  $f$  is guaranteed to have an absolute maximum and absolute minimum. Furthermore, these must occur either at a critical point or on the boundary of  $D$ .*

**Example 10.7.8 Absolute values on a triangle.** Let  $f(x, y) = x + y - xy$ . Find the absolute maxima and minima of  $f$  on the triangle with vertices  $(0, 0)$ ,  $(0, 2)$  and  $(4, 0)$ .

**Solution.** We'll start off by finding any critical points. Since  $\nabla f = \langle 1 - y, 1 - x \rangle$ , we see that the only critical point is  $(1, 1)$ . Now we need to see what happens to  $f(x, y)$  on the edges of the triangle, and to do that we need to find equations for the edges. Let  $l_1$  denote the edge from  $(0, 0)$  to  $(0, 2)$ ,  $l_2$  the edge from  $(0, 2)$  to  $(4, 0)$  and  $l_3$  the edge from  $(0, 0)$  to  $(4, 0)$ . Then  $l_1$  is given by the equation  $x = 0$ ,  $l_2$  is given by  $y = -\frac{1}{2}(x - 4)$  and  $l_3$  is given by  $y = 0$ . So on  $l_1$  we have  $f(x, y) = f(0, y) = y$ , on  $l_2$  we have

$$f(x, y) = f\left(x, -\frac{1}{2}(x - 4)\right) = \frac{1}{2} \left[ \left(x - \frac{3}{2}\right)^2 + \frac{7}{4} \right]$$

and on  $l_3$  we have  $f(x, y) = f(x, 0) = x$ . We can then set up the following table:

edge	$f(x, y)$	max	min
$l_1$	$y$	2	0
$l_2$	$\frac{1}{2} \left[ \left(x - \frac{3}{2}\right)^2 + \frac{7}{4} \right]$	$\frac{1}{2} \left[ \left(4 - \frac{3}{2}\right)^2 + \frac{7}{4} \right] = 4$	$\frac{7}{8}$
$l_3$	$x$	4	0

**Table 10.7.9:** Absolute values on the boundary

Finally, at the critical point  $(1, 1)$  we have  $f(1, 1) = 1$ . So putting this all together, we can say that the absolute maximum of  $f(x, y)$  on this triangle is 4, and this is reached at  $(4, 0)$ . Similarly, the absolute minimum is 0, and this is reached at the origin.

Note that in [Example 10.7.8](#), we don't actually determine if the critical point is a local maximum or minimum. We just compare the value  $f$  reaches at the critical point with its maximum and minimum values on the boundary.

**Example 10.7.10 Absolute maxima and minima on a circular sector.** Let  $f(x, y) = xy^2$  and let

$$D = \{(x, y) : x, y \geq 0, x^2 + y^2 \leq 3\}$$

Find the absolute maxima and minima of  $f$  on  $D$ .

<sup>1</sup>This means that  $D$  is finite in size and that it includes its boundary.

**Solution.** First we'll find the critical point(s). If we compute  $\nabla f$ , we get  $\nabla f = \langle y^2, 2xy \rangle$ . This is 0 if and only if  $y = 0$ , and if  $y = 0$  then this forces  $f(x, y) = f(x, 0) = 0$ .

Now we need to determine the maximum and minimum values of  $f(x, y)$  on the boundary of  $D$ . Let  $C_1$  denote the line segment from  $(0, 0)$  to  $(\sqrt{3}, 0)$ ,  $C_2$  the circular arc from  $(\sqrt{3}, 0)$  to  $(0, \sqrt{3})$  and  $C_3$  the line segment from  $(0, \sqrt{3})$  to  $(0, 0)$ . Then on  $C_1$  we have  $f(x, y) = 0$ , on  $C_2$  we have

$$f(x, y) = xy^2 = x(3 - x^2)$$

and on  $C_3$  we have  $f(x, y) = 0$ . So the smallest value that  $f$  takes on the boundary is 0, and the largest value is 2. So the absolute maximum is 2 and the absolute minimum is 0.

SUGGESTED PROBLEMS: 1, 3, 13, 23

## 10.8 Lagrange Multipliers

[Theorem 10.7.4](#) is useful for finding local maxima and minima of a function  $f(x, y)$ , but it's often the case where we have an additional restriction on the variables  $x$  and  $y$ , which we interpret as an equality  $g(x, y) = k$  for some constant  $k$ . Geometrically, this reduces down to finding the maximum and minimum values of  $f(x, y)$  on the curve  $g(x, y) = k$ . In these cases, [Theorem 10.7.4](#) is not as useful because there's no guarantee that the critical points of  $f$  will lie on the curve  $g(x, y) = k$ . For this scenario, we must use the method of **Lagrange multipliers**.

**Algorithm 10.8.1 Method of Lagrange Multipliers.** *Suppose that  $f$  and  $g$  are both differentiable functions, and suppose that  $\nabla g \neq \mathbf{0}$  on the curve  $g(x, y) = k$ . To find the extreme values of  $f(x, y)$  subject to the constraint  $g(x, y) = k$ , perform the following steps:*

1. Find all points  $(x_0, y_0)$  such that  $\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0)$ , where  $\lambda$  is a constant we call the **Lagrange multiplier**.
2. Compute  $f$  at all points from the first step and compare values to find any maxima or minima.

**Example 10.8.2 Maximizing volume.** A lidless box may be made from  $12 \text{ m}^2$  of material. What dimensions give the maximum volume?

**Solution.** If we denote the dimensions of the box by  $x, y, z$ , then the function we need to maximize is  $V(x, y, z) = xyz$ . However, not all dimensions are valid since we're only allowed to use  $12 \text{ m}^2$  of material. This means that the surface area of the box, which is given by  $A(x, y, z) = xy + 2xz + 2yz$ , must be equal to 12. So our constraint is that  $A(x, y, z) = 12$ , and because we have this constraint we'll try to solve this using [Algorithm 10.8.1](#).

To begin, we need to solve  $\nabla V = \lambda \nabla A$ . Equivalently, we need to solve the system

$$V_x = \lambda A_x$$

$$V_y = \lambda A_y$$

$$V_z = \lambda A_z$$

which turns into

$$yz = \lambda(y + 2z)$$

$$\begin{aligned}xz &= \lambda(x + 2z) \\xy &= \lambda(2x + 2y)\end{aligned}$$

Now, we don't actually need to find a value for  $\lambda$ . We just need values for  $x, y, z$ . If we multiply these equations by  $x, y$  and  $z$  respectively, then we get

$$\begin{aligned}xyz &= \lambda(xy + 2xz) \\xyz &= \lambda(xy + 2yz) \\xyz &= \lambda(2xz + 2yz)\end{aligned}$$

or just

$$\lambda(xy + 2xz) = \lambda(xy + 2yz) = \lambda(2xz + 2yz)$$

We can also divide by  $\lambda$ , since there's no way  $\lambda$  could equal 0 and satisfy the above equations (since  $x, y, z$  are all positive). This gives us

$$xy + 2xz = xy + 2yz = 2xz + 2yz \quad (10.8.1)$$

The first two expressions in (10.8.1) simplify to  $x = y$ , while the first and third reduce to  $x = 2z$ . So

$$x = y = 2z.$$

But these variables also need to satisfy  $A(x, y, z) = 12$ :

$$\begin{aligned}12 &= xy + 2xz + 2yz \\&= 4z^2 + 4z^2 + 4z^2 \\&= 12z^2\end{aligned}$$

Therefore  $z = \pm 1$ , so we take  $z = 1$  (since we're dealing with dimensions). This also means that  $x = y = 2(1) = 2$ . So the point  $(2, 2, 1)$  is therefore our candidate for the dimensions that maximize volume subject to the constraint  $A = 12$ .

In order to actually verify that  $V(2, 2, 1) = 4$  is actually the largest possible volume we can obtain, we need to show that it's actually a maximum and not a minimum. This means we need to find another point  $(x, y, z)$  that also satisfies  $A = 12$ , and show that it gives a smaller value for  $V$ . One such point is just  $(-2, -2, -1)$ , and  $V(-2, -2, -1) = -4 < V(2, 2, 1)$ . So 4 is actually a maximum value and not a minimum value.

The points we'll obtain using Lagrange multipliers will be either maxima or minima, but the method itself doesn't always tell us which is which. As in the last example, sometimes it's up to us to find that on our own.

**Example 10.8.3 Extreme values on a paraboloid.** Find the absolute maxima and minima for  $f(x, y) = x^2 + 2y^2$  over this disk  $x^2 + y^2 \leq 1$ .

**Solution.** First, recall from Section 10.7 that the absolute maxima and minima of a differentiable function over a bounded set occur at either a critical point or somewhere on the boundary. So first, we need to find if  $f(x, y)$  has any critical points inside the disk. If we solve  $f_x = f_y = 0$ , then we quickly get that  $(0, 0)$  is a critical point, and hence a possible maximum or minimum value. The only other place we need to check is the boundary  $x^2 + y^2 = 1$ , and this is something we can use Lagrange multipliers for.

Set  $g(x, y) = x^2 + y^2$ . Then we need to find the extreme values of  $f$  subject to the constraint  $g = 1$ . The system of equations we need to solve is then

$$2x = \lambda(2x) \quad \Rightarrow \quad x = \lambda x$$



$$\begin{aligned} 4y &= \lambda(2y) & \Rightarrow 2y &= \lambda y \\ x^2 + y^2 &= 1 \end{aligned}$$

The first equation is true if  $\lambda = 1$ , but then this forces  $y = 0$  in the second equation. If we then plug  $y = 0$  into the constraint, this forces  $x = \pm 1$ . So the case  $\lambda = 1$  gives us two points to test:  $(-1, 0)$  and  $(1, 0)$ . On the other hand, the second equation is true if  $\lambda = 2$ . This then forces  $x = 0$  and  $y = \pm 1$ , which gives us the points  $(0, -1)$  and  $(0, 1)$ . Finally, if  $\lambda \neq 1, 2$ , then this forces  $x = y = 0$ . But there's no way to satisfy the constraint if  $x = y = 0$ , so we discard this possibility.

So we have five points that we need to check, which we arrange in the following table:

$(x, y)$	$f(x, y)$
$(0, 0)$	0
$(-1, 0)$	1
$(1, 0)$	1
$(0, -1)$	2
$(0, 1)$	2

**Table 10.8.4:** Absolute values on the disk

So the absolute minimum is 0, which occurs at the origin. The absolute maximum is 2, which occurs at the points  $(0, \pm 1)$ .

**Example 10.8.5 Maximizing the volume of a prism within an ellipsoid.** A rectangular prism, centered at the origin and with sides parallel to the coordinate planes, is inscribed within the ellipsoid  $x^2 + \frac{y^2}{4} + \frac{z^2}{4} = 1$ . Find the dimensions of the prism that maximize the volume.

**Solution.** Let  $(x, y, z)$  denote the corner of the prism located in the first octant. Then we want to maximize  $V(x, y, z) = (2x)(2y)(2z)$  subject to the constraint  $g(x, y, z) = x^2 + \frac{y^2}{4} + \frac{z^2}{4} = 1$ . So we'll use the method of Lagrange multipliers again to get the system

$$\begin{aligned} 8yz &= 2\lambda x \\ 8xz &= \frac{\lambda}{2}y \\ 8xy &= \frac{\lambda}{2}z \\ x^2 + \frac{y^2}{4} + \frac{z^2}{4} &= 1 \end{aligned}$$

We can assume that  $x, y, z$  are all positive (since we're trying to find the *maximum* volume), so we can go ahead and solve each equation for  $\lambda$  to get

$$\lambda = \frac{4yz}{x} = \frac{16xz}{y} = \frac{16xy}{z}$$

Setting the first two fractions equal and simplifying gives  $y^2 = 4x^2$ . Similarly, the second and third fractions give  $y^2 = z^2$ . So

$$y^2 = z^2 = 4x^2.$$

Now we plug this into our constraint to get

$$1 = x^2 + \frac{y^2}{4} + \frac{z^2}{4} = 1$$

$$= 3x^2$$

so  $x = \pm\sqrt{\frac{1}{3}}$ . Since we're assuming that  $(x, y, z)$  lies in the first octant, this means that  $x > 0$ , along with  $y$  and  $z$ . So

$$\begin{aligned}x &= \sqrt{\frac{1}{3}} \\y &= 2\sqrt{\frac{1}{3}} \\z &= 2\sqrt{\frac{1}{3}}\end{aligned}$$

To check that these values actually maximize  $V(x, y, z)$ , we can compare them with the point  $(1, 0, 0)$ , which also satisfies the constraint  $g = 1$ . Since  $V(1, 0, 0) = 0$ , this means that the dimensions that maximize the volume are given by

$$\frac{2}{\sqrt{3}} \times \frac{4}{\sqrt{3}} \times \frac{4}{\sqrt{3}}$$

SUGGESTED PROBLEMS: 1-11 odd, 17, 19

# Chapter 11

## Multiple Integrals

From Calculus I, we know how to find the area under the graph  $y = f(x)$  from  $x = a$  to  $x = b$ : it's just

$$\int_a^b f(x) dx.$$

What we want to do now is to take the notion of integration and extend it to higher dimensions. As a starting motivation, we'd like to develop the concept of an integral of a function  $f(x, y)$  of two independent variables in order to find the volume under the surface  $z = f(x, y)$ .

### 11.1 Double Integrals over Rectangles

We want to try to find the volume under the surface  $z = f(x, y)$  and above some rectangle  $R$  in the  $xy$ -plane. If  $f(x, y)$  is a constant function, then this is easy: the resulting surface and the  $xy$ -plane then form a rectangular prism. If  $f(x, y)$  is more complicated, then we might not have a nice volume formula to use. However, we can steal an idea from Calculus I and try to approximate the volume by using simpler shapes; in this case, rectangular prisms.

#### 11.1.1 Riemann Sums

Given a continuous function  $f(x, y)$  defined over the rectangle

$$R = \{(x, y) : a \leq x \leq b, c \leq y \leq d\}$$

to approximate the volume under  $f$  and above  $R$  we'll divide the rectangle up into smaller rectangles. In particular, let's divide the intervals  $[a, b]$  and  $[c, d]$  into subintervals  $[x_i, x_{i+1}]$ ,  $[y_j, y_{j+1}]$ , where

$$a = x_0 < x_1 < x_2 < \cdots < x_m = b$$

$$c = y_0 < y_1 < y_2 < \cdots < y_n = d$$

$$\Delta x = x_{i+1} - x_i$$

$$\Delta y = y_{j+1} - y_j$$

If we let  $R_{ij}$  denote the subrectangle given by

$$R_{ij} = \{(x, y) : x_i \leq x \leq x_{i+1}, y_j \leq y \leq y_{j+1}\}$$

then the volume under  $f(x, y)$  and above  $R_{ij}$  is about equal to

$$f(x_i, y_j) \Delta x \Delta y$$

Hence the volume under  $f(x, y)$  and above  $R$  should be about equal to

$$\sum_{i=0}^m \sum_{j=0}^n f(x_i, y_j) \Delta x \Delta y$$

If we replace  $\Delta x \Delta y$  with  $\Delta A$  and send  $\Delta A \rightarrow 0$ , this approximation becomes exact. This gives us the following definition.

**Definition 11.1.1 Double Integral Over a Rectangle.** Let  $f(x, y)$  be a function defined over the rectangle  $R$ . Then the **double integral** of  $f(x, y)$  over  $R$  is defined to be

$$\iint_R f(x, y) dA = \lim_{\Delta A \rightarrow 0} \sum_{i=0}^m \sum_{j=0}^n f(x_i, y_j) \Delta A$$

If the limit exists, we say that  $f$  is **integrable** over  $R$ .

The geometric intuition behind the double integral is that it represents a volume under a surface. Negative values for a double integral indicate that a surface tends to lie below the  $xy$ -plane, while positive values indicate that the surface tends to lie above the  $xy$ -plane. Furthermore, continuous functions are integrable over any rectangle  $R$ .

**Example 11.1.2 Double integrals by volume.** Let  $f(x, y) = 5 - x$  and let

$$R = \{(x, y) : 0 \leq x \leq 5, 0 \leq y \leq 3\}$$

Find  $\iint_R f(x, y) dA$ .

**Solution.** If we graph  $f(x, y)$ , we get a triangular cylinder running along the  $y$  axis. The volume of this cylinder over  $R$  is just the area of the triangular "base" times the "height," or just

$$\frac{1}{2} 5 \cdot 5 \cdot 3 = \frac{75}{2}$$

So

$$\iint_R (5 - x) dA = \frac{75}{2}$$

### 11.1.2 Fubini's Theorem

Now that we have a definition for the double integral, we want to find a better way of computing it. Thankfully, we do have such a method: Fubini's Theorem.

**Theorem 11.1.3 Fubini's Theorem.** Let  $f(x, y)$  be continuous on the rectangle

$$R = \{(x, y) : a \leq x \leq b, c \leq y \leq d\}$$

Then

$$\iint_R f(x, y) dA = \int_a^b \left[ \int_c^d f(x, y) dy \right] dx = \int_c^d \left[ \int_a^b f(x, y) dx \right] dy$$

**Example 11.1.4 Double integrals by Fubini's Theorem.** Let  $f(x, y) = 5 - x$  and let

$$R = \{(x, y) : 0 \leq x \leq 5, 0 \leq y \leq 3\}$$

Find  $\iint_R f(x, y) dA$ .

**Solution.** We found this previously in [Example 11.1.2](#), so we'll try it again using Fubini's Theorem and see if we get the same answer. By [Theorem 11.1.3](#), we have

$$\begin{aligned}\iint_R f(x, y) dA &= \int_0^3 \left[ \int_0^5 (5-x) dx \right] dy \\ &= \int_0^3 \left[ 5x - \frac{x^2}{2} \right]_{x=0}^5 dy \\ &= \int_0^3 \frac{25}{2} dy \\ &= \frac{75}{2}\end{aligned}$$

The double integral, like the Calculus I integral, derivative and partial derivatives, is *linear*. This means that you can break it up over addition and pull constants outside of it. In other words, if  $f$  and  $g$  are integrable and if  $c$  is a constant, then

$$\begin{aligned}\iint_R (f(x, y) + g(x, y)) dA &= \iint_R f(x, y) dA + \iint_R g(x, y) dA \\ \iint_R cf(x, y) dA &= c \iint_R f(x, y) dA\end{aligned}$$

**Example 11.1.5 Double integral of a logarithm.** Find

$$\int_1^3 \int_1^3 \frac{\ln(xy)}{xy} dy dx$$

**Solution.** First, note that

$$\int_1^3 \int_1^3 \frac{\ln(xy)}{xy} dy dx = \int_1^3 \int_1^3 \frac{\ln(x)}{xy} dy dx + \int_1^3 \int_1^3 \frac{\ln(y)}{xy} dy dx$$

To find the first double integral on the right, first we integrate with respect to  $y$  to get

$$\int_1^3 \int_1^3 \frac{\ln(x)}{xy} dy dx = \int_1^3 \frac{\ln x}{x} \ln y \Big|_{y=1}^3 dx = \ln 3 \int_1^3 \frac{\ln x}{x} dx$$

Now we can use  $u$ -sub with  $u = \ln x$  and  $du = \frac{1}{x} dx$  to get

$$\begin{aligned}\ln 3 \int_1^3 \frac{\ln(x)}{x} dx &= \ln 3 \frac{u^2}{2} \Big|_{u=0}^{\ln 3} \\ &= \frac{1}{2} (\ln 3)^3\end{aligned}$$

The same exact process shows that  $\int_1^3 \frac{\ln y}{xy} dy dx$  must also equal  $\frac{1}{2} (\ln 3)^3$ . Therefore

$$\int_1^3 \int_1^3 \frac{\ln(xy)}{xy} dy dx = (\ln 3)^3$$

SUGGESTED PROBLEMS: 7, 9, 15, 19

## 11.2 Double Integrals over General Regions

Section 11.1 shows how to define the double integral over a rectangle  $R$ . Now we want to extend it to more general regions. We'll be skipping an issue with how to actually define such an integral in terms of Riemann sums, but the main idea is to take a function defined over some general region  $D$  and then extend it to cover a rectangular region  $R$ . Then we can just use Definition 11.1.1 to define this new integral as well.

Our primary tool for computing double integrals will still be Theorem 11.1.3. The main difficulty we'll encounter (aside from the usual integration issues) is how to properly set up limits for  $x$  and  $y$  for some region  $D$ .

**Example 11.2.1 Integrating over a region defined by a parabola.** Let  $f(x, y) = \frac{y}{x^5+1}$  and let  $D$  be the region under the parabola  $y = x^2$ , above the  $x$ -axis and between  $x = 0$  and  $x = 1$ . Find  $\iint_D f(x, y) dA$ .

**Solution.** The first step is to determine limits for  $x$  and  $y$  that describe this region. One possible description is the following:

$$D = \{(x, y) : 0 \leq y \leq x^2, 0 \leq x \leq 1\}$$

If we use this, we can write

$$\begin{aligned} \iint_D \frac{y}{x^5+1} dA &= \int_0^1 \int_0^{x^2} \frac{y}{x^5+1} dy dx \\ &= \frac{1}{2} \int_0^1 \frac{1}{x^5+1} y^2 \Big|_{y=0}^{x^2} dx \\ &= \frac{1}{2} \int_0^1 \frac{x^4}{x^5+1} dx \\ &= \frac{1}{10} \int_1^2 \frac{1}{u} du \\ &= \frac{1}{10} \ln 2 \end{aligned}$$

We also could have described  $D$  as the set of all points

$$D = \{(x, y) : \sqrt{y} \leq x \leq 1, 0 \leq y \leq 1\}$$

If we do this instead, we get

$$\iint_D \frac{y}{x^5+1} dA = \int_0^1 \int_{\sqrt{y}}^1 \frac{y}{x^5+1} dx dy$$

By Theorem 11.1.3, this is guaranteed to equal  $\frac{1}{10} \ln 2$ . However, the actual process of computing the double integral using this choice for the limits is much more difficult, since we would need to integrate  $\frac{1}{x^5+1}$  with respect to  $x$ .

There are several important things we can take away from the above example:

- Much of the difficulty in computing double integrals lies in finding limits that describe the region you're integrating over. In general, it's a good idea to sketch the region you're integrating over in order to figure out what your limits should be.
- There can be multiple ways to describe a single region. This leads to multiple ways of setting up limits for double integrals over this region. If the integrand is continuous, then Theorem 11.1.3 guarantees that these different methods will all produce the same value.

- It's important to choose limits in such a way as to make computing the double integral more manageable.
- When finding limits for a double integral, the outermost limits *must be constant* since the double integral must eventually simplify to a constant. In other words, we don't really have a notion of an "indefinite" double integral.

**Example 11.2.2 Reversing the order of integration.** Compute  $\int_0^{\sqrt{\pi}} \int_y^{\sqrt{\pi}} \cos x^2 \, dx \, dy$ .

**Solution.** If we try to integrate with respect to  $x$  right away, then we're stuck:  $\cos x^2$  has no elementary antiderivative, which for us means that we can only integrate it using power series. However, it is continuous, so [Theorem 11.1.3](#) tells us that if we can integrate it with respect to  $y$  first and then  $x$  (i.e. replace  $dx \, dy$  with  $dy \, dx$ ), we'll still get the same answer. And integrating  $\cos x^2$  with respect to  $y$  is actually very easy. This is called **reversing the order of integration**.

In order to change the order of integration, we need to change the limits. The region we're integrating over is the region

$$D = \{(x, y) : y \leq x \leq \sqrt{\pi}, 0 \leq y \leq \sqrt{\pi}\}$$

If we sketch this, we see that this is just the region below the line  $y = x$ , above the  $x$ -axis and bounded from  $x = 0$  to  $x = \sqrt{\pi}$ . So we can also write

$$D = \{(x, y) : 0 \leq y \leq x, 0 \leq x \leq \sqrt{\pi}\}$$

Therefore

$$\begin{aligned} \int_0^{\sqrt{\pi}} \int_y^{\sqrt{\pi}} \cos x^2 \, dx \, dy &= \int_0^{\sqrt{\pi}} \int_0^x \cos x^2 \, dy \, dx \\ &= \int_0^{\sqrt{\pi}} y \cos x^2 \Big|_{y=0}^x \, dx \\ &= \int_0^{\sqrt{\pi}} x \cos x^2 \, dx \\ &= \frac{1}{2} \int_0^{\pi} \cos u \, du \\ &= 0 \end{aligned}$$

Sometimes, the order of integration can be chosen to make the limits of integration easier to set up.

**Example 11.2.3 Integrating over a Triangle.** Integrate  $f(x, y) = x^2 + y^2$  over the triangle  $T$  with vertices  $(0, 0)$ ,  $(1, 1)$  and  $(0, 2)$ .

**Solution.** If we wanted to integrate with respect to  $x$  first, we would have to break our double integral up into two different double integrals. This is because the limits for  $x$  change halfway up the triangle. So it'd be better for us to integrate with respect to  $y$  first. Since

$$T = \{(x, y) : x \leq y \leq 2 - x, 0 \leq x \leq 1\}$$

we have

$$\iint_T (x^2 + y^2) \, dA = \int_0^1 \int_x^{2-x} (x^2 + y^2) \, dy \, dx$$

$$\begin{aligned}
&= \int_0^1 \left[ x^2 y + \frac{y^3}{3} \right]_{y=x}^{2-x} dx \\
&= \int_0^1 \left[ x^2(2-2x) + \frac{1}{3}[(2-x)^3 - x^3] \right] dx \\
&= \left[ \frac{2x^3}{3} - \frac{1}{2}x^4 - \frac{(2-x)^4}{12} - \frac{x^4}{4} \right]_0^1 \\
&= \frac{2}{3} - \frac{1}{2} - \frac{1}{12} - \frac{1}{4} + \frac{16}{12} \\
&= \frac{7}{6}
\end{aligned}$$

Besides finding volumes, we can also use double integrals to find areas. If  $D$  is some region in the  $xy$ -plane, then  $\iint_D dA$  represents the volume under the surface  $z = 1$  and above  $D$ . This is a solid with the constant height of 1, so the volume should be equal to the area of  $D$  times 1. That is,  $\iint_D dA$  is equal to the area of  $D$ .

SUGGESTED PROBLEMS: 1, 5, 11, 13, 17, 21

### 11.3 Double Integrals in Polar Coordinates

Recall that the double integral was defined by first setting up a rectangular grid. The reason we used a rectangular grid was because we were working in Cartesian coordinates, so this made the most sense. If we're dealing with a circular region of integration, then using Cartesian coordinates is very awkward. However, polar coordinates from [Section 8.3](#) work very nicely with circular regions. So we want to find out how to set up double integrals using polar coordinates.

If we're given a function  $f(x, y)$ , then it's not too hard to convert this to the polar form  $f(r, \theta)$ . Just replace  $x$  with  $r \cos \theta$  and  $y$  with  $r \sin \theta$ . The tricky part with setting up double integrals in polar coordinates is how to deal with the **area element**  $dA$ , which in Cartesian coordinates is just  $dx dy$  or  $dy dx$ . To figure out what  $dA$  should be in polar coordinates, i.e. in terms of  $r$  and  $\theta$ , consider the following "polar rectangle":

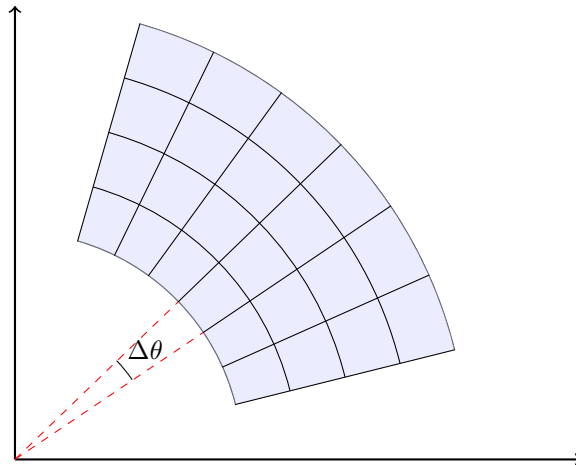


Figure 11.3.1: A polar grid.

Let  $\Delta A$  represent the area of one of these sectors. If we let  $r$  denote the distance from the origin to one sector,  $\Delta r$  the length of a sector and  $\Delta \theta$  the



angle spanned by a sector, then we can say that

$$\begin{aligned}\Delta A &\approx \frac{1}{2}[r \cos(\Delta\theta)][r \sin(\Delta\theta)] - \frac{1}{2}[(r - \Delta r) \cos(\Delta\theta)][(r - \Delta r) \sin(\Delta\theta)] \\ &= r\Delta r \cos(\Delta\theta) \sin(\Delta\theta) - \frac{1}{2}(\Delta r)^2 \cos(\Delta\theta) \sin(\Delta\theta)\end{aligned}$$

If we assume that  $\Delta r$  and  $\Delta\theta$  are both small (which means the polar grid in Figure 11.3.1 is very fine), then

$$\begin{aligned}\cos(\Delta\theta) &\approx 1 \\ \sin(\Delta\theta) &\approx \Delta\theta \\ (\Delta r)^2 &\approx 0\end{aligned}$$

So

$$\Delta A \approx r\Delta r\Delta\theta.$$

As  $\Delta r$  and  $\Delta\theta$  approach 0, this becomes more exact, and we get

$$dA = r \, dr \, d\theta.$$

**Theorem 11.3.2 Double Integrals in Polar Coordinates.** *Let  $f(x, y)$  be a continuous function. Then*

$$\iint_D f(x, y) \, dA = \iint_D f(r \cos \theta, r \sin \theta) r \, dr \, d\theta$$

*and limits are chosen using polar coordinates.*

**Example 11.3.3 Integrating over a circular sector.** Find

$$\int_{-1}^0 \int_{-\sqrt{1-x^2}}^0 \frac{2}{1 + \sqrt{x^2 + y^2}} \, dy \, dx$$

**Solution.** If we sketch the region of integration, we see that it is the part of the unit circle in the third quadrant. So we'll switch to polar coordinates to solve this integral:

$$\begin{aligned}\int_{-1}^0 \int_{-\sqrt{1-x^2}}^0 \frac{2}{1 + \sqrt{x^2 + y^2}} \, dy \, dx &= \int_{\pi}^{\frac{3\pi}{2}} \int_0^1 \frac{2}{1 + r} r \, dr \, d\theta \\ &= 2 \int_{\pi}^{\frac{3\pi}{2}} \frac{r}{1 + r} \, dr \, d\theta \\ &= 2 \int_{\pi}^{\frac{3\pi}{2}} \left[ r - \ln(1 + r) \right]_{r=0}^1 \, d\theta \\ &= \pi(1 - \ln 2)\end{aligned}$$

Polar coordinates may also be used, surprisingly, to evaluate the **Gaussian integral**  $\int_{-\infty}^{\infty} e^{-x^2} \, dx$ .

**Theorem 11.3.4 The Gaussian Integral.** *We have*

$$\int_{-\infty}^{\infty} e^{-x^2} \, dx = \sqrt{\pi}$$

*Proof.* First, let  $I = \int_{-\infty}^{\infty} e^{-x^2} dx$ . We'll show that  $I^2 = \pi$ . We have

$$\begin{aligned} I^2 &= \left( \int_{-\infty}^{\infty} e^{-x^2} dx \right) \left( \int_{-\infty}^{\infty} e^{-y^2} dy \right) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy \\ &= \int_0^{2\pi} \int_0^{\infty} e^{-r^2} r dr d\theta \\ &= \int_0^{2\pi} \lim_{b \rightarrow \infty} \left[ \frac{-1}{2} e^{-r^2} \right]_{r=0}^b d\theta \\ &= \int_0^{2\pi} \frac{1}{2} d\theta \\ &= \pi \end{aligned}$$

Since  $I^2 = \pi$ , this gives

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$$

□

**Example 11.3.5 Volume of a sphere.** Find the volume of a sphere of radius  $\rho$ .

**Solution.** First, we can center the sphere at the origin without loss of generality. Such a sphere is given by  $x^2 + y^2 + z^2 = \rho^2$ . If we solve for  $z$ , we get

$$z = \pm \sqrt{\rho^2 - x^2 - y^2}$$

Let  $D$  denote the disk of radius  $\rho$  in the  $xy$ -plane centered at the origin. Then the volume of the sphere is

$$\begin{aligned} 2 \iint_D \sqrt{\rho^2 - x^2 - y^2} dA &= 2 \int_0^{2\pi} \int_0^{\rho} \sqrt{\rho^2 - r^2} r dr d\theta \\ &= - \int_0^{2\pi} \int_{\rho^2}^0 \sqrt{u} du d\theta \\ &= \frac{2}{3} \int_0^{2\pi} u^{3/2} \Big|_{u=\rho^2}^0 d\theta \\ &= \frac{4\pi}{3} \rho^3 \end{aligned}$$

SUGGESTED PROBLEMS: 5, 13, 23

## 11.4 Applications of Double Integrals

### 11.4.1 Mass

Consider a thin plate in the  $xy$ -plane, say the region  $R$ . If the density of the plate  $\rho(x, y)$  at  $(x, y)$  is constant, say  $\rho(x, y) = C$ , then the mass of the plate is just the density times the area. On the other hand, if the density varies then this becomes more complicated, and we must use double integrals. In particular, the mass of a plate contained in the region  $R$  in the  $xy$ -plane with density  $\rho(x, y)$  is given by

$$\iint_R \rho(x, y) dA$$

**Example 11.4.1 Mass of a triangular plate.** Find the mass of the plate contained in the triangular region  $D$  bounded by lines  $y = x$ ,  $y = 6$  and  $2x + y = 6$ , given that the density is  $\rho(x, y) = x^2$ .

**Solution.** The mass is

$$\begin{aligned}\iint_D x^2 A &= \int_2^6 \int_{\frac{6-y}{2}}^y x^2 dx dy \\ &= \frac{1}{3} \int_2^6 \left[ \frac{y^3 - (6-y)^3}{8} \right] dy\end{aligned}$$

The **moments** of the mass contained in  $D$  and with density  $\rho(x, y)$  are defined as follows:

$$\begin{aligned}M_x &= \iint_D y \rho(x, y) dA \\ M_y &= \iint_D x \rho(x, y) dA\end{aligned}$$

If we let  $m$  denote the total mass, then we also define the **center of mass** (or **centroid**) to be the point  $(\bar{x}, \bar{y})$ , where

$$\begin{aligned}\bar{x} &= \frac{M_y}{m} \\ \bar{y} &= \frac{M_x}{m}\end{aligned}$$

**Example 11.4.2 Center of mass of an annulus.** Find the center of mass of the plate  $D$  outside of the circle  $x^2 + y^2 = 1$  and inside the circle  $x^2 + y^2 = 4$ , with density  $\rho(x, y) = \frac{1}{\sqrt{x^2 + y^2}}$ .

**Solution.** The mass is given by

$$\begin{aligned}\iint_D \frac{1}{\sqrt{x^2 + y^2}} dA &= \int_0^{2\pi} \int_1^2 \frac{r}{r} dr d\theta \\ &= 2\pi\end{aligned}$$

The moments are

$$\begin{aligned}M_x &= \iint_D \frac{y}{\sqrt{x^2 + y^2}} dA \\ &= \int_0^{2\pi} \int_1^2 \frac{r \sin \theta}{r} r dr d\theta \\ &= 0\end{aligned}$$

and likewise  $M_y = 0$ . So the center of mass is  $(0, 0)$ .

SUGGESTED PROBLEMS: 3, 5, 13

## 11.5 Triple Integrals

In  $\mathbb{R}$ , we have

$$\int_a^b f(x) dx = \lim_{\Delta x \rightarrow 0} \sum_{i=0}^n f(x_i) \Delta x.$$

This represents the area under  $y = f(x)$  and over  $[a, b]$ . Furthermore,  $\int_a^b dx$  gives the length of  $[a, b]$ . In  $\mathbb{R}^2$ , we have

$$\int_D f(x, y) dA = \lim_{\Delta A \rightarrow 0} \sum_{i=0}^m \sum_{j=0}^n f(x_i, y_j) \Delta A,$$

This represents the volume under  $z = f(x, y)$  and above the region  $D$ , where  $\Delta A = \Delta x \Delta y$ . Furthermore,  $\iint_D dA$  gives the area of  $D$ .

We can extend all of this to  $\mathbb{R}^3$  by introducing the concept of the **triple integral**.

**Definition 11.5.1 Triple Integrals over a Rectangle.** Let  $f(x, y, z)$  be defined on some region  $D$  in  $\mathbb{R}^3$ . Then the triple integral of  $f$  over  $D$  is given by

$$\iiint_D f(x, y, z) dV = \lim_{\Delta V \rightarrow 0} \sum_{i=0}^l \sum_{j=0}^m \sum_{k=0}^n f(x_i, y_j, z_k) \Delta V$$

where  $\Delta V = \Delta x \Delta y \Delta z$ . If the limit exists, we say that  $f$  is **integrable** on  $D$ .

For a double integral in rectangular coordinates, we have  $dA = dx dy$  or  $dy dx$ . Similarly, for a triple integral in rectangular coordinates we have six different choices for  $dV$ :

$dx dy dz$	$dx dz dy$
$dy dx dz$	$dy dz dx$
$dz dx dy$	$dz dy dx$

**Table 11.5.2:** Absolute values on the disk

Just as we can view  $dx$  as an infinitesimal length and  $dA$  as an infinitesimal area,  $dV$  represents an infinitesimal volume. Our main use for [Definition 11.5.1](#) will be to recognize a triple integral "in the wild," but we won't actually use it to compute integrals. For this purpose, we still use Fubini's Theorem.

**Theorem 11.5.3 Fubini's Theorem for Triple Integrals.** Suppose  $f(x, y, z)$  is a continuous function on the closed and bounded region  $D$  in  $\mathbb{R}^3$ . Then  $\iiint_D f(x, y, z) dV$  can be computed as an iterated integral, and the answer does not depend on the choice of  $dV$ .

**Example 11.5.4 A triple integral over a rectangular prism.** Compute  $\iiint_D xyz e^{-x^2-y^2} dV$ , where

$$D = \{(x, y) : 0 \leq x \leq \sqrt{\ln 2}, 0 \leq y \leq \sqrt{\ln 4}, 0 \leq z \leq 1\}$$

**Solution.** We'll integrate using  $dV = dx dy dz$ . Then we have

$$\begin{aligned} \iiint_D xyz e^{-x^2-y^2} dV &= \int_0^1 \int_0^{\sqrt{\ln 4}} \int_0^{\sqrt{\ln 2}} xyz e^{-x^2-y^2} dx dy dz \\ &= \int_0^1 \int_0^{\sqrt{\ln 4}} \left[ -\frac{yz e^{-x^2-y^2}}{2} \right]_{x=0}^{\sqrt{\ln 2}} dy dz \\ &= -\frac{1}{2} \int_0^1 \int_0^{\sqrt{\ln 4}} yz (e^{-\ln 2 - y^2} - e^{-y^2}) dy dz \\ &= -\frac{e^{-\ln 2} - 1}{2} \int_0^1 \int_0^{\sqrt{\ln 4}} yz e^{-y^2} dy dz \end{aligned}$$

$$\begin{aligned}
&= \frac{-\frac{1}{2}}{4} \int_0^1 \left[ ze^{-y^2} \right]_{y=0}^{\sqrt{\ln 4}} dz \\
&= \frac{\frac{3}{4}}{8} \int_0^1 z dz \\
&= \frac{\frac{3}{32}}{2} \\
&= \frac{3}{64}
\end{aligned}$$

An unfortunate side effect of increasing the dimension for our integral is that we lose a little bit of geometric intuition. For instance, [Example 11.5.4](#) is indeed calculating a "volume," but the volume in question is for a four dimensional region (the graph of  $f(x, y, z)$  over the rectangular prism). We can only really visualize the "base" of this region, which served as our region of integration in [Example 11.5.4](#). Even so, the triple integral can still tell us important things about functions of three variables.

**Example 11.5.5 Finding an average value.** Find the average value of the function  $f(x, y, z) = x y z e^{-x^2 - y^2}$  over the region  $D$  given in [Example 11.5.4](#).

**Solution.** First, let  $\text{Vol}(D)$  denote the volume of  $D$ . Then the average value of  $f$  over  $D$  is just

$$\frac{\iiint_D x y z e^{-x^2 - y^2} dV}{\text{Vol}(D)} = \frac{\frac{3}{64}}{\sqrt{\ln 2} \ln 4} \approx .05$$

We can also compute triple integrals over more general regions.

**Example 11.5.6 Volume using triple integrals.** Find the volume of the region bounded by the cylinder  $y = x^2$  and the planes  $z = 3 - y$  and  $z = 0$ .

**Solution.** If we let  $D$  denote this region, then its volume is given by  $\iiint_D dV$ . The volume is then

$$\begin{aligned}
\iiint_D dV &= \int_{-\sqrt{3}}^{\sqrt{3}} \int_{x^2}^3 \int_0^{3-y} dz dy dx \\
&= \int_{-\sqrt{3}}^{\sqrt{3}} \int_{x^2}^3 (3 - y) dy dx \\
&= \int_{-\sqrt{3}}^{\sqrt{3}} \left[ 3y - \frac{y^2}{2} \right]_{y=x^2}^3 dx \\
&= \int_{-\sqrt{3}}^{\sqrt{3}} \left[ 9 - \frac{9}{2} - 3x^2 + \frac{x^2}{2} \right] dx \\
&= \frac{24\sqrt{3}}{2}
\end{aligned}$$

When setting up limits for triple integrals, say using  $dV = dx dy dz$ , then the limits on the innermost integral are typically functions of  $y$  and  $z$ , the limits on the middle integral are functions of  $z$  and the limits on the outermost integral are constant. We can also change the order of integration to make an integral more tractable.

**Example 11.5.7 Changing the order of integration.** Compute  $\int_1^4 \int_z^{4z} \int_0^{\pi^2} \frac{\sin \sqrt{yz}}{x^{3/2}} dy dx dz$ .

**Solution.** This looks awful to integrate with respect to  $y$  first, so we'll try changing the order of integration.  $x$  looks easiest, so let's try using  $dV =$

$dx dy dz$  instead. If we sketch the region, we see that the limits are actually the same, except we just need to swap the middle and innermost integrals. So

$$\begin{aligned}
 \int_1^4 \int_z^{4z} \int_0^{\pi^2} \frac{\sin \sqrt{yz}}{x^{3/2}} dy dx dz &= \int_1^4 \int_0^{\pi^2} \int_z^{4z} x^{-3/2} \sin \sqrt{yz} dx dy dz \\
 &= \int_1^4 \int_0^{\pi^2} \left[ -2x^{-1/2} \sin \sqrt{yz} \right]_{x=z}^{4z} dy dz \\
 &= \int_1^4 \int_0^{\pi^2} z^{-1/2} \sin \sqrt{yz} dy dz \\
 &= \int_0^{\pi^2} \int_1^4 z^{-1/2} \sin \sqrt{yz} dz dy \\
 &= \int_0^{\pi^2} \int_{\sqrt{y}}^{2\sqrt{y}} \frac{2}{\sqrt{y}} \sin u du dy \\
 &= - \int_0^{\pi^2} \left[ \frac{2}{\sqrt{y}} \cos u \right]_{u=\sqrt{y}}^{2\sqrt{y}} dy \\
 &= \int_0^{\pi^2} \frac{2}{\sqrt{y}} (\cos \sqrt{y} - \cos 2\sqrt{y}) dy \\
 &= 4 \int_0^{\pi} (\cos v - \cos 2v) dv \\
 &= 0
 \end{aligned}$$

SUGGESTED PROBLEMS: 5, 11, 13, 19, 25, 29

## 11.6 Cylindrical Coordinates

In [Section 8.3](#), we saw that introducing a new coordinate system made certain double integrals much easier to work with. The same idea works with triple integrals. The first such system we'll look at is **cylindrical coordinates**, which are useful for computing triple integrals over cylindrical regions. To convert Cartesian coordinates  $(x, y, z)$  into cylindrical coordinates  $(r, \theta, z)$ , simply replace  $x, y$  with polar coordinates, and use  $dV = r dr d\theta dz$ . We leave  $z$  alone.

**Example 11.6.1 Computing a triple integral over a cylinder.** Let  $D$  be the cylinder in  $\mathbb{R}^3$  with height 2 and base given by the circle of radius 4 centered at the origin, restricted to the first and second octants. Let  $f(x, y, z) = \frac{z}{1+x^2+y^2}$ . Compute  $\iiint_D f(x, y, z) dV$ .

**Solution.** Since we're dealing with a cylinder, we'll switch to cylindrical coordinates  $r dr d\theta dz$ :

$$\begin{aligned}
 \iiint_D f(x, y, z) dV &= \int_0^2 \int_0^{\pi} \int_0^4 \frac{r}{1+r^2} dr d\theta dz \\
 &= \pi \ln 17
 \end{aligned}$$

# Appendix A

## GNU Free Documentation License

Version 1.3, 3 November 2008

Copyright © 2000, 2001, 2002, 2007, 2008 Free Software Foundation, Inc.  
<<http://www.fsf.org/>>

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

**0. PREAMBLE** The purpose of this License is to make a manual, textbook, or other functional and useful document “free” in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondly, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of “copyleft”, which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

**1. APPLICABILITY AND DEFINITIONS** This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The “Document”, below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as “you”. You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A “Modified Version” of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A “Secondary Section” is a named appendix or a front-matter section of

the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

The "publisher" means any person or entity that distributes copies of the Document to the public.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards



disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

**2. VERBATIM COPYING** You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

**3. COPYING IN QUANTITY** If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

**4. MODIFICATIONS** You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

- A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You

may use the same title as a previous version if the original publisher of that version gives permission.

- B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.
- C. State on the Title page the name of the publisher of the Modified Version, as the publisher.
- D. Preserve all the copyright notices of the Document.
- E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
- F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
- G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
- H. Include an unaltered copy of this License.
- I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.
- J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.
- K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.
- L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
- M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.
- N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.
- O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties — for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

**5. COMBINING DOCUMENTS** You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

**6. COLLECTIONS OF DOCUMENTS** You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

**7. AGGREGATION WITH INDEPENDENT WORKS** A compilation of the Document or its derivatives with other separate and independent

documents or works, in or on a volume of a storage or distribution medium, is called an “aggregate” if the copyright resulting from the compilation is not used to limit the legal rights of the compilation’s users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document’s Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

**8. TRANSLATION** Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled “Acknowledgements”, “Dedications”, or “History”, the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

**9. TERMINATION** You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it.

**10. FUTURE REVISIONS OF THIS LICENSE** The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See <http://www.gnu.org/copyleft/>.

Each version of the License is given a distinguishing version number. If

the Document specifies that a particular numbered version of this License “or any later version” applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy’s public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

**11. RELICENSING** “Massive Multiauthor Collaboration Site” (or “MMC Site”) means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A “Massive Multiauthor Collaboration” (or “MMC”) contained in the site means any set of copyrightable works thus published on the MMC site.

“CC-BY-SA” means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

“Incorporate” means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is “eligible for relicensing” if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.

**ADDENDUM: How to use this License for your documents** To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

```
Copyright (C)  YEAR  YOUR NAME.
Permission is granted to copy, distribute and/or modify this document
under the terms of the GNU Free Documentation License, Version 1.3
or any later version published by the Free Software Foundation;
with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.
A copy of the license is included in the section entitled "GNU
Free Documentation License".
```

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the “with... Texts.” line with this:

```
with the Invariant Sections being LIST THEIR TITLES, with the
Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.
```

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.



# Index

- Antiderivatives, 50
- chain rule
  - one independent variable, 144
  - two independent variables, 145
- continuous functions
  - definition, 5
- cross product
  - definition, 117
  - magnitude, 118
  - orthogonality, 117
  - properties, 118
- cylinders
  - definition, 125
- derivative
  - $n^{\text{th}}$ -order derivative, 17
  - definition, 13
  - derivative function, 14
  - differentiation formulas
    - chain rule, 21
    - constant multiple rule, 19
    - power rule, 18
    - product rule, 20
    - sum rule, 19
  - linear approximation, 25
- derivatives
  - differentiation formulas
    - sine and cosine, 19
- directional derivative
  - definition, 146
  - formula, 146
- distance formula
  - three dimensions, 108
- dot product
  - definition, 113
  - geometry, 114
  - properties, 113
- double integrals
  - over a rectangle, 156
  - polar coordinates, 161
- exponential functions
  - derivatives, 32
  - properties, 27
- Extreme Value Theorem, 42
- Fermat's Theorem, 42
- Fubini's Theorem
  - triple integrals, 164
- functions
  - absolute extrema, 41
  - differentiable, 13
  - inverse functions
    - derivatives, 30
  - invertible functions, 29
  - local extrema, 41
- functions of several variables
  - contour plots, 138
  - two variables, 137
    - continuity, 139
    - limits, 138
- gradient
  - direction of greatest increase/decrease, 148
- integral convergence tests
  - direct comparison test, 84
  - limit comparison test, 85
- integration techniques
  - integration by parts, 69
- inverse trigonometric functions
  - inverse sine, 35
- L'Hospital's Rule, 38
- Lagrange multipliers, 151
- limit
  - function of a single variable, 2
- linearization, 143
- lines
  - parametric equations in  $\mathbb{R}^3$ , 122
- logarithm functions
  - derivatives, 33
  - properties, 31

- maxima and minima
  - absolute, [149](#)
  - Extreme Value Theorem, [150](#)
  - local, [148](#)
  - second derivative test, [149](#)
- Mean Value Theorem, [44](#)
- partial derivatives, [140](#)
- partial fractions
  - algorithm, [78](#)
- planes
  - vector and scalar equations, [123](#)
- polar coordinates
  - arc length, [106](#)
  - area, [104](#)
- scalar triple product, [119](#)
- tangent planes, [143](#)
- torque, [119](#)
- triple integrals
  - rectangular coordinates, [164](#)
- vector function
  - components of acceleration, [135](#)
- vector functions
  - binormal vector, [134](#)
  - curvature, [132](#)
    - alternative formula, [133](#)
  - derivative rules, [130](#)
  - derivatives, [129](#)
  - unit normal vector, [134](#)
- vectors
  - properties of vector addition and scalar multiplication, [110](#)
  - standard basis, [111](#)
  - unit vectors, [111](#)
  - vector projection, [116](#)



This book was authored in MathBook XML.