# MAT133
# Course Notes

Tyler Holden
Mathematics and Computational Sciences
University of Toronto Mississauga
tyler.holden@utoronto.ca

# Contents

# 1 Introduction

We begin with a rapid review of preliminary concepts. This material will form the foundation of what's to come, so it is essential that you have a mastery of these concepts.

## 1.1 Sets and notation

A *set* is any collection of well-defined and distinct objects. By this we mean that you can put as many things as you like into a set, so long as they are concrete and all different. We often surround the elements of a set by curly braces $\{,\}$, for example

$$\{1,2,3,...\}, \qquad \{\text{cat}, \text{dog}, \text{bird}\}, \qquad \{\ddot\smile, \ddot\frown\}, \qquad \{\heartsuit, \clubsuit, \diamondsuit, \spadesuit\}.$$

We can put anything we want into a set[1] so long as the object is a well-defined thing (for example, we cannot consider the set of all objects which I think are interesting. What objects are in this set? It is ambiguous), and all the elements of the set are distinct (so the object $\{1,1,2\}$ is not a set, because the element 1 appears multiple times).

We use the symbol '$\in$' (read as 'in') to talk about when an element is in a set; for example, $1 \in \{1,2,3\}$ but $\ddot\frown \notin \{\text{dog}, \text{cat}\}$. We can also talk about subsets, which are collections of items in a set and indicated with a '$\subseteq$' sign. For example,

$$\{2,4,6\} \subseteq \{1,2,3,4,5,6\}$$

since every element on the left-hand-side is also present in the right-hand-side.

Since sets can have many objects within them, it is often impractical to list them all explicitly. Instead, we might use *set-builder* notation, which allows to say "the set of all things which satisfy some property." For example,

$$\{x : x > 0\}$$

is read as "the set of all $x$ such that $x$ is greater than 0," while

$$\{\text{month} : \text{month ends in 'ber'}\} = \{\text{September, October, November, December}\},$$

is the set of all months for which the end of the name of the month ends in 'ber'. Another way we could say this, is to let $M$ be the set of all months, and consider the set

$$\{x \in M : x \text{ ends in 'ber'}\}.$$

More interesting than months are sets of numbers, since they tend to be quite big. Some sets that we will be involved with a lot are as follows:

- The **empty set** $\emptyset$, which has nothing inside of it.
- The **naturals**[2] $\mathbb{N} = \{0,1,2,3,\ldots\}$,

[1] This is not true, but it's quite a tangent to explain why.
[2] Some mathematicians do not believe that 0 is a natural number.

- The **integers** $\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}$,
- The **rationals** $\mathbb{Q} = \{p/q : p, q \in \mathbb{Z}, q \neq 0\}$,
- The **reals** $\mathbb{R}$ (the set of all infinite decimal expansion).

The real numbers are our focus in this course. Roughly speaking, the real numbers consist of all possible decimal expansions, and are denoted $\mathbb{R}$. For example, $\pi$, $\sqrt{2}$, and $-17.37\overline{125}$, are real numbers.

To discuss collections of numbers, we often use *sets*. Sets are often written in the form

$$S = \{x : x \text{ satisfies some property}\}.$$

An example of this are intervals in $\mathbb{R}$: If $a, b$ are real numbers with $a < b$, we reite

$$\begin{aligned} (a,b) &= \{x : a < x < b\}, & [a,b) &= \{x : a \leq x < b\}, \\ (a,b] &= \{x : a < x \leq b\}, & [a,b] &= \{x : a \leq x \leq b\}. \end{aligned}$$

In particular, a parenthesis means that the endpoint is *not* included in the interval, while a square bracket indicates that the endpoint is contained in the interval. We say that the interval $(a, b)$ is an *open* interval and $[a, b]$ is a *closed* interval. The intervals $(a, b]$ and $[a, b)$ may be referred to as either half open or half closed. When we wish to indicate that $x$ is simply less than or larger than a number, we include a $\pm\infty$ sign in the appropriate spot. For example,

$$\begin{aligned} (-\infty, a) &= \{x : x < a\} & (-\infty, a] &= \{x : x \leq a\} \\ (b, \infty) &= \{x : x > b\} & [b, \infty) &= \{x : x \geq b\} \end{aligned}$$

Note that the infinity sign is always used in conjunction with an open bracket. If you are familiar with the notion of unions and intersections, you can use these to combine intervals in a convenient way. We use the $\in$ symbol to indicate when an element is in a set. For example, $2 \in (-3, 3)$, but $4 \notin [0, 1]$.

In our studies we will come across $\mathbb{R}^n$, the collection of $n$-tuples of real numbers. For example, $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$ are the couples of real numbers – often visualized as the two dimensional plane – while $\mathbb{R}^3 = \{(x, y, z) : x, y, z \in \mathbb{R}\}$ are the triples of real numbers.

### 1.1.1 Operations on Sets

**Union and Intersection:** Let $S$ be a set and choose two sets $A, B \subseteq S$. We define the *union* of $A$ and $B$ to be

$$A \cup B = \{x \in S : x \in A \text{ or } x \in B\}$$

and the *intersection* of $A$ and $B$ to be

$$A \cap B = \{x \in S : x \in A \text{ and } x \in B\}.$$

**Example 1.1**

Determine the union and intersection of the following two sets:

$$A = \{x \in \mathbb{R} : x > 1\}, \qquad B = \{x \in \mathbb{R} : -1 < x < 2\}.$$

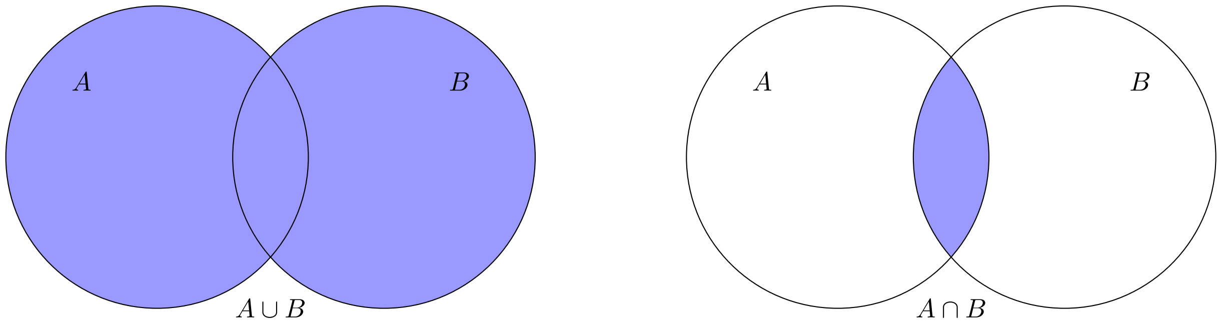



Figure 1.1: Left: The union of two sets is the collection of all elements which are in both (though remember that elements of sets are distinct, so we do not permit duplicates). Right: The intersection of two sets consists of all elements which are common to both sets.

*Solution.* By definition, one has

$$\begin{aligned} A \cup B &= \{x \in \mathbb{R} : x \in A \text{ or } x \in B\} = \{x \in \mathbb{R} : x > 1 \text{ or } -1 < x < 2\} \\ &= \{x \in \mathbb{R} : x > -1\}, \\ A \cap B &= \{x \in \mathbb{R} : x \in A \text{ and } x \in B\} = \{x \in \mathbb{R} : x > 1 \text{ and } -1 < x < 2\} \\ &= \{x \in \mathbb{R} : 1 < x < 2\}. \end{aligned}$$

■

**Complement** If $A \subseteq S$ then the *complement* of $A$ with respect to $S$ is all elements which are not in $A$; that is,

$$A^c = \{x \in S : x \notin A\}.$$

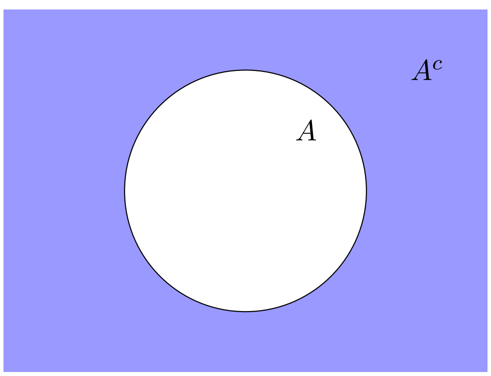



Figure 1.2: The complement of a set $A$ with respect to $S$ is the set of all elements which are in $S$ but not in $A$.

**Example 1.2**

Let $A$ and $B$ be defined as in Example 1.1. Find the complements of $A$ and $B$ in $\mathbb{R}$.

*Solution.* $A = \{x \in \mathbb{R} : x > 1\}$, and consists of all those numbers which are strictly larger than one. Its complement are those elements which are not strictly larger than one; namely, those which are less than 1. Hence

$$A^c = \{x \in \mathbb{R} : x \leq 1\} = (-\infty, 1].$$

Similarly, $B = (-1, 2)$ will have a complement consisting of those numbers less than $-1$ and greater than 2, or

$$B^c = (-\infty, -1] \cup [2, \infty).$$ ■

**Set Difference:** Given two sets $A$ and $B$, their difference is

$$A \setminus B = \{x \in A : x \notin B\}.$$

For example, let

$$A = \{1, 2\}, \quad B = \{2, 3, 4\}, \quad \text{and} \quad C = \{1, 5, 6\}.$$

The set differences are as follows:

$$\begin{array}{cc} A \setminus B = \{1\} & A \setminus C = \{2\} \\ B \setminus A = \{3, 4\} & B \setminus C = \{2, 3, 4\} \\ C \setminus A = \{5, 6\} & C \setminus B = \{1, 5, 6\} \end{array}$$

**Cartesian Product** The Cartesian product of two sets $A$ and $B$ is the collection of ordered pairs, one from $A$ and one from $B$; namely,

$$A \times B = \{(a, b) : a \in A, b \in B\}.$$

For example, if $C = \{H, T\}$ and $D = \{1, 2, 3, 4, 5, 6\}$, then

$$C \times D = \{(H, 1), (H, 2), (H, 3), (H, 4), (H, 5), (H, 6), (T, 1), (T, 2), (T, 3), (T, 4), (T, 5), (T, 6)\}.$$

Higher dimensional spaces can be constructed using the Cartesian product. For example, we know that we can represent the plane $\mathbb{R}^2$ as an ordered pair of points $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$, while three dimensional space is an ordered triple $\mathbb{R}^3 = \{(x, y, z) : x, y, z \in \mathbb{R}\}$. In this sense, we see that $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$, and motivates the more general definition of $\mathbb{R}^n$ as an ordered $n$-tuple

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{-times}}.$$

## 1.2 Functions

We think of a function as a machine which eats a number and produces another number. It is important that a function only produce a single output for each input. For example, the function $f(x) = x^2$ takes in an input $x$ and produces the output $x^2$.

| Input | Output |
|---|---|
| $-2$ | $4$ |
| $0$ | $0$ |
| $\sqrt{2}$ | $2$ |
| $\pi$ | $\pi^2$ |

A function has a *domain* and a *range*. The domain is the set of all things which can be put into the function, while the range is the set of all things which come out of the function.

**Example 1.3**

Determine the domain and range of each of the following functions:

1. $f(x) = \dfrac{1}{x}$,
2. $g(x) = \sqrt{x-2}$,
3. $h(x) = (x-1)^2 - 3$,
4. $r(x) = \dfrac{1}{\sqrt{(x-1)(x+1)}}$.

*Solution.*

1. We may divide by every number except 0, hence the domain of this function is $(-\infty, 0) \cup (0, \infty)$. For the range, we notice that $1/x$ can never be zero, since if so then $1/x = 0$ implies that $1 = 0$, and this cannot be true. Hence the range is also $(-\infty, 0) \cup (0, \infty)$.

2. Since we may not take the square root of a negative number, we require that $x - 2 \geq 0$ or rather, $x \geq 2$. Hence $f$ has domain $[2, \infty)$. On the other hand, the square root function is always non-negative, with minimum occurring at $x = 2$, showing that the range of $g$ is $[0, \infty)$.

3. We have no restrictions on what numbers can be input into $h$, so the domain of $h$ is $\mathbb{R}$. The range requires a bit more thought. Notice that the value of $(x-1)^2$ is always non-negative, regardless of the input, so $(x-1)^2 - 3 \geq -3$. This is in fact the range $[-3, \infty)$.

4. Since we cannot divide by zero, the points $x = \pm 1$ cannot be in the domain of $r$. Similarly, we cannot take the square root of a negative number. We can determine where $(x+1)(x-1) > 0$ with the following table:

| | $x < -1$ | $-1 < x < 1$ | $x > 1$ |
|---|---|---|---|
| $x - 1$ | $-$ | $-$ | $+$ |
| $x + 1$ | $-$ | $+$ | $+$ |
| $(x-1)(x+1)$ | $+$ | $-$ | $+$ |

so that $(x-1)(x+1)$ is positive when $x < -1$ and $x > 1$; that is, on the interval $(-\infty, 1) \cup (1, \infty)$. Hence this is the domain of $r$. The range is much tougher! Try it on your own. ■

A useful way to visualize functions is in terms of a graph. This is the collection of points in the $xy$-plane with coordinate $(x, f(x))$; that is, the collection of points in the $xy$-plane whose horizontal distance is the $x$-value and whose vertical distance is $f$, seen in Figure 1.3. The graph of a function gives us the ability to discern qualitative properties of a function by visualizing its behaviour.

Given a curve in the $xy$-plane, there is a simple way of determining whether that curve is the graph of a function, called the *vertical line test*. Since a function must send each input $x$ to a *unique* output $f(x)$, this tells us that for each $x \in \mathbb{R}$, there can be at most one point of the graph lying above it. Given a fixed $x_0$, we may determine the corresponding $f(x_0)$ by drawing a vertical line through $x_0$ and seeing where this line intersects the given curve. If there is any point on a graph where a vertical line intersects the curve twice, the curve cannot correspond to a function.

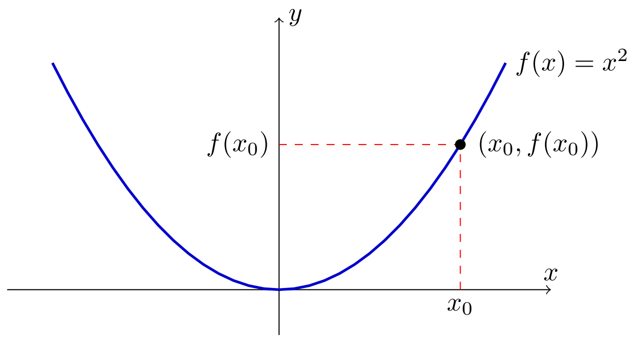



Figure 1.3: The graph of the function $f(x) = x^2$. The dark blue line is the graph of $f$, embedded in the plane. In the plane, its coordinates are just $(x, f(x))$.

**Example 1.4**

Consider the curves given in Figure 1.4. Determine which are given by functions, and which fail to be functions.

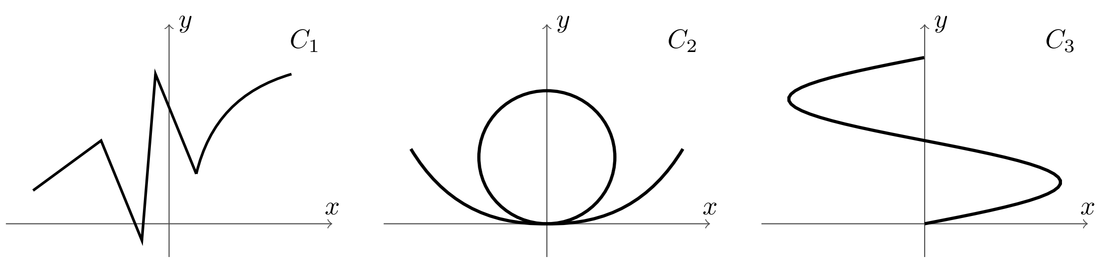



Figure 1.4: A collection of curves. Which of these are given by functions and which cannot possibly be given by functions?

*Solution.* The curve corresponding to $C_1$ is not particularly appealing, but is nonetheless given by a function. Regardless of where we choose to draw a vertical line, it will intersect the graph at only one point.

The curves $C_2$ and $C_3$ cannot possibly be the graphs of functions, as they fail the vertical line test. While there are many points at which the test fails, perhaps the most obvious place is the $y$-axis. This axis is indeed a vertical line and intersects each of $C_2$ and $C_3$ in 2 and 3 points respectively. ■

### 1.2.1 Operations on Functions

Functions may be added and multiplied in a *pointwise* manner. For example, if $f$ and $g$ are functions, then

$$(f + g)(x) = f(x) + g(x) \quad \text{and} \quad (fg)(x) = f(x)g(x).$$

When the range of one function includes into the domain of another, we may combine the two to form a new function $g \circ f$, known as the *composition*:

$$(g \circ f)(s) = g(f(s)).$$

You can think of this as chaining together two functions. Note that not all functions can be composed, and the restriction that the range of $f$ be part of the domain of $g$ is essential. For example, the function $f(x) = -x^2 - 1$ has range $(-\infty, -1]$ and the function $g(x) = \sqrt{x}$ has domain $[0, \infty)$. The composition $(g \circ f)(x) = \sqrt{-x^2 - 1}$ does not make sense, since any input of $x$ would require that we take the square root of a negative number, which we cannot do.

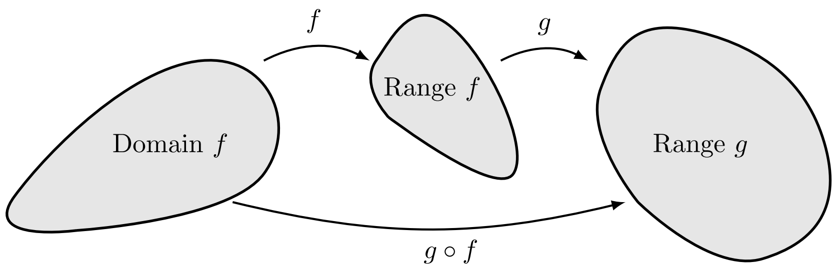



Figure 1.5: Given two functions $f(x)$ and $g(x)$, their composition $g \circ f$.

**Example 1.5**

Consider the three functions $f(x) = 4x, g(x) = 1/(x^2 + 1)$ and $h(x) = x + 7$. Compute $f \circ g, f \circ h, g \circ h$ and compare these to $g \circ f, h \circ f$ and $h \circ g$.

*Solution.* Computing compositions can be as easy as substituting one function into the argument of another. Hence

$$(f \circ g)(x) = f(g(x)) = f\left(\frac{1}{x^2+1}\right) = \frac{4}{x^2+1}.$$

Continuing in this fashion for all other examples, we find that

$$\begin{aligned} (f \circ g)(x) &= \frac{4}{x^2+1} & (g \circ f) &= \frac{1}{16x^2+1} \\ (f \circ h)(x) &= 4x + 28 & (h \circ f) &= 4x + 7 \\ (g \circ h)(x) &= \frac{1}{x^2 + 14x + 50} & (h \circ g) &= \frac{7x^2+8}{x^2+1}. \end{aligned}$$

Note that in general, the order of the compositions matters. ■

### 1.2.2 Symmetries

Functions which exhibit symmetric about an axis often have nice properties which make them simple to study. Here we will talk about what it means for a function to be even or odd.

**Definition 1.6**

A function $f$ is said to be *even* if $f(-x) = f(x)$ for all $x \in \mathbb{R}$. In turn, $f$ is said to be *odd* if $f(-x) = -f(x)$ for all $x \in \mathbb{R}$.

Let's take a moment to determine what is happening. Assume first that $f(x)$ is an even function, so that $f(-x) = f(x)$. This means that for a fixed value $x_0$, the height of the graph of $f(x)$ is the same at both $x_0$ and $-x_0$, hence even functions are symmetric about reflections in the $y$-axis. On the other hand, if $f(x)$ is an odd function so that $f(-x) = -f(x)$, then the height of the graph at $-x_0$ is the same as at $x_0$ but now negative, hence odd functions are symmetric about rotations of $180°$. Figure 1.6 gives examples of even and odd functions.

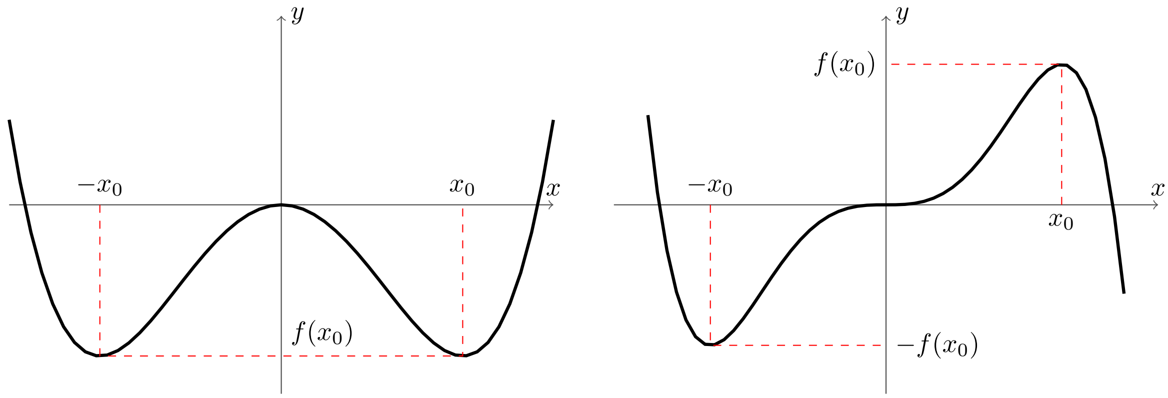



Figure 1.6: Examples of even (left) and odd (right) functions. Notice that the left figure exhibits reflectional symmetry about the $y$-axis, while the right figure exhibits a rotational symmetry of $180°$ (or equivalently, in both the $x$- and $y$-axes).

**Example 1.7**

Determine whether the following functions are even, odd, or neither:

$$f(x) = x^2, \quad g(x) = x^3, \quad h(x) = f(x) + g(x).$$

*Solution.* To determine whether a function has any of these symmetries, substitute $-x$ into its argument and see if you can relate it to the original function. For $f$ we have

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

implying that $f$ is even. For $g$ we have

$$g(-x) = (-x)^3 = -x^3 = -g(x)$$

implying that $g$ is odd. Finally, for $h$ we have

$$h(-x) = f(-x) + g(-x) = f(x) - g(x).$$

However, there is no natural way to relate $f - g$ to $f + g$ by using only a single minus sign. Hence $h$ is neither even nor odd. ∎

### 1.2.3 Roots

Mathematically, the number 0 is an interesting if sometimes troublesome number. You are likely familiar with the fact that $0 \times a = 0$ and $0 + a = a$ for any value of $a$, and that division by 0 is strictly prohibited. Hence it is unsurprising that we give special consideration to when a function takes on this value.

**Definition 1.8**

If $f$ is a function, we say that $\alpha \in \mathbb{R}$ is a *root* of $f$ if $f(\alpha) = 0$. Geometrically, roots correspond to the places at which the graph of a function passes through the $x$-axis.

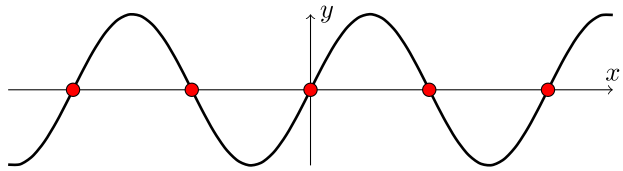



Figure 1.7: The roots of a function $f$ correspond to those instances where it crosses the $x$-axis (red circles).

**Example 1.9**

Find the roots of the functions

$$f_1(x) = x - 5, \qquad f_2(x) = 0, \qquad f_3(x) = \frac{1}{x}.$$

*Solution.* We begin by looking at $f_1$. We're looking for those values $\alpha$ for which $f_1(\alpha) = \alpha - 5 = 0$, which can be solved to find that $\alpha = 5$. This is the only possible root of $f_1$. It is easy to see that given any function of the form $g(x) = x - r$, the root of $g$ will be $r$.

For $f_2$, we want the collection of $\alpha$ satisfying $f_2(\alpha) = 0$. Since $f_2$ is just the function which sends everything to zero, it turns out that every real number is a root of $f_2$. This turns out to be clear when we realize that the graph of $f_2$ is just the $x$-axis itself.

Finally, for $f_3$ we want $\alpha$ such that $f_3(\alpha) = 1/\alpha = 0$. In order to solve this equation for $\alpha$, we would need to take a reciprocal of both sides, but this would require us to divide by zero! Hence $1/\alpha = 0$ has no solutions, implying that $f_3$ has not roots. Again, try plotting $f_3$ and this will become obvious. ■

### 1.2.4 Piecewise Functions

Piecewise functions are described by gluing together two functions to form a new one. For example, if $g$ and $h$ are two functions and $a \in \mathbb{R}$, we may define

$$f(x) = \begin{cases} g(x) & x \leq a \\ h(x) & x > a \end{cases}.$$

This means that if $x \leq a$ then $f(x) = g(x)$ and if $x > a$ then $f(x) = h(x)$.

**Example 1.10**

Graph the function $f(x) = \begin{cases} x^2 & x \leq 0 \\ 3 & x > 0 \end{cases}$.

*Solution.* The graphs of $y = x^2$, $y = 3$, and $f(x)$ are given in Figure 1.8. It is our hope that this illustrates the idea of a piecewise function in terms of cutting and pasting; namely, we cut the graphs of $x^2$ and 3 at the line $x = 0$ and then re-attach them in the way described by $f(x)$. ■

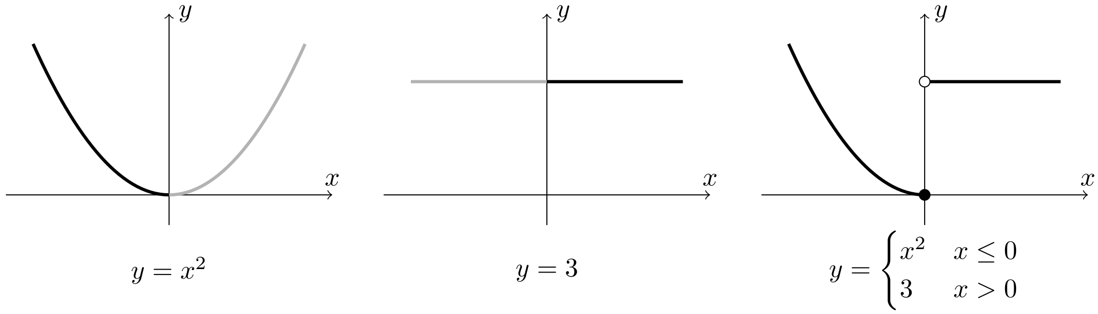



Figure 1.8: A piecewise function is a way of gluing two functions together to form a new function. This figure illustrates how we have taken the functions $y = x^2$ and $y = 3$, cut each along the line $x = 0$, and then glued them together to get the function $f$. Notice the exaggerated hole in $f$, used to indicate that the value 3 is not actually attained at $x = 0$.

While our examples above utilized two functions, there is no limit on the number of functions which we may splice together, so long as that number is finite. For example, the following piecewise function has four components:

$$f(x) = \begin{cases} 2 - x^2 & x < 1 \\ 0 & x = 1 \\ 4 - x & 1 < x < 2 \\ 1/x & x > 2 \end{cases}.$$

Try graphing this piecewise function.

### 1.2.5 Inverse Functions

The word "inverse" has many different meanings depending on the context in which it is used. For example, what if we were to ask the student to find the inverse of the number 2? What does this mean? To what are we taking the inverse? To properly understand this, we need to understand the following: Given a binary operator (an operator which takes in two things and produces a single thing in return, such as addition and multiplication), we say that a number id is the *identity* of that operator if operating against it does nothing to the input. For example, in the case of addition, the operator will satisfy $x + \mathrm{id}_+ = x$ for all possible $x$; for example,

$$2 + \mathrm{id}_+ = 2, \qquad -5 + \mathrm{id}_+ = -5.$$

Our experience tells us that $\mathrm{id}_+ = 0$. Similarly, for multiplication the identity $\mathrm{id}_\times$ will satisfy $x \times \mathrm{id}_\times = x$ for all $x$; for example,

$$3 \times \mathrm{id}_\times = 3, \qquad \pi \times \mathrm{id}_\times = \pi.$$

Again our experience tells us that $\mathrm{id}_\times = 1$. We thus say that 0 is the additive identity and 1 is the multiplicative identity. We say that the *inverse of* $x$ is an element which, when paired against $x$, gives the identity. Hence the additive inverse of 2 is the number $y$ such that $2 + y = \mathrm{id}_+ = 0$, or rather $-2$. In general, the additive inverse of $n$ is $-n$, and this always exists! For multiplication, it is not too hard to convince ourselves that the multiplicative inverse of $x$ is $1/x$; for example, $2 \times (1/2) = 1 = \mathrm{id}_\times$. Notice that there is no multiplicative inverse for the number 0, so in this case the inverse does not always exist.

Function composition $f \circ g$ is another example of a binary operator. What is the identity for this operation? Well, we would like a function $\mathrm{id}_\circ$ such that

$$\begin{aligned} f(\mathrm{id}_\circ(x)) &= f(x) \\ &= \mathrm{id}_\circ(f(x)). \end{aligned}$$

If we this about this for a moment, the identity function is the function $\mathrm{id}_\circ(x) = x$, the function which does nothing to the argument! Now what is the inverse of a function? The inverse of a function $f$ is a <u>function</u> $f^{-1}$ such that $f \circ f^{-1} = f^{-1} \circ f = \mathrm{id}_\circ$.

To compute the inverse of $y = f(x)$, notice that by applying $f^{-1}$ to both sides we get

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

Hence by switching $x$ and $y$ and solving for $y$, we get $y = f^{-1}(x)$.

**Example 1.11**

Determine the inverse of the function $y = f(x) = (x-1)/(x+1)$.

*Solution.* As recommended above, we interchange $y$ and $x$ and solve for $y$, so we get

$$\begin{aligned} x = \frac{y-1}{y+1} \quad &\Leftrightarrow \quad (y+1)x = y - 1 \\ &\Leftrightarrow \quad yx - y = -(x+1) \\ &\Leftrightarrow \quad y(x-1) = -(x+1) \\ &\Leftrightarrow \quad y = \frac{x+1}{1-x} \end{aligned}$$

So $f^{-1}(x) = (x+1)/(1-x)$. Indeed we can check this by composing $f \circ f^{-1}$ and $f^{-1} \circ f$ to find that

$$\begin{aligned} f(f^{-1}(x)) &= \frac{\frac{x+1}{1-x} - 1}{\frac{x+1}{1-x} + 1} \\ &= \frac{\frac{x+1-(1-x)}{1-x}}{\frac{x+1+1-x}{1-x}} = \frac{2x}{2} \\ &= x \end{aligned}$$

and the other direction is left as an exercise. ■

Note that not all functions are invertible. For example, the function $f(x) = x^2$ is not invertible in general. It is tempting to say that $g(x) = \sqrt{x}$ is the inverse to $f$, but this is not the case. Indeed, while we do have that

$$(f \circ g) = \left(\sqrt{x}\right)^2 = x,$$

the opposite composition gives

$$(g \circ f)(x) = \sqrt{x^2} = |x|,$$

which is not the identity function. To test whether a function can be inverted, it must satisfy the *horizontal line test*; that is, every horizontal line must intersect the graph of $f$ in at most one place.

**Exercise**

Determine the inverses of each of the following functions. What special property do $f, g$, and $h$ all share?

$$f(x) = \frac{1}{x}, \qquad g(x) = 1 - x, \qquad h(x) = \frac{x}{x-1}.$$

## 1.3 Polynomials and Rational Functions

Polynomials are the collection of all objects of the form

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0$$

for some natural number $n > 0$ and real numbers $a_0, a_1, \ldots, a_n$. We say that the *degree* of a polynomial $p$, written $\deg(p)$, is the highest power whose coefficient is non-zero. For example, the following functions are polynomials:

$$p(x) = 3x^4 + 8x^3 - 2x, \qquad q(x) = 39x^{66} - 5x^2 + 1$$

and $\deg(p) = 4$ while $\deg(q) = 66$. Some degrees occur so frequently that they even have special names:

| Degree | Name |
|---|---|
| 1 | Linear |
| 2 | Quadratic |
| 3 | Cubic |
| 4 | Quartic |
| 5 | Quintic |

Factoring polynomials is the process by which we reverse the act of multiplying a polynomial; that is, we would like to write a single polynomial as a product of polynomials with strictly smaller degree.

There are some very easy factorization which involve simply removing a power of $x$. If there is no constant term (the coefficient of $x^0$ is 0), then we may remove at least one power of $x$ from the polynomial. For example,

$$x^4 + x^2 = x^2(x^2 + 1), \qquad x^5 + x^4 + x = x(x^4 + x^3 + 1).$$

In general, factoring polynomials with constant terms can be difficult. For most purposes however, we may limit ourselves to factoring quadratic polynomials. Given a quadratic polynomial of the form $x^2 + ax + b$, the trick is to try and find two numbers $p, q$ such that $a = p + q$ and $b = pq$. This is because

$$(x + p)(x + q) = x^2 + (p + q)x + pq.$$

**Example 1.12**

Factor the following polynomials:

$$x^2 + 2x + 1, \qquad 3x^2 + 15x + 18, \qquad x^2 - 1, \qquad x^3 - x^2 - 2x.$$

*Solution.* We begin with $x^2 + 2x + 1$. To factor this, try thinking of two numbers $p, q$ such that $p + q = 2$ and $pq = 1$. Hopefully, the choice $p = 1, q = 1$ springs to our minds and we guess $(x + 1)(x + 1) = x^2 + 2x + 1$. A quick check verifies that this is the case.

For $3x^2 + 15x + 18$ we are not quite in the situation described above as the coefficient in front of $x^2$ is not 1. However, we may first factor out a 3 to get $3x^2 + 15x + 18 = 3(x^2 + 5x + 6)$. Now we would like to find $p, q$ such that $p + q = 5$ and $pq = 6$. The choice $p = 2$ and $q = 3$ jumps to mind, and a quick calculation verifies that $(x + 2)(x + 3) = x^2 + 5x + 6$. Thus

$$3x^2 + 15x + 18 = 3(x + 2)(x + 3).$$

The polynomial $x^2 - 1$ looks tricky: what do we do if we have no $x$ term? Instead of panicking, let's try our usual technique; that is, find $p, q$ such that $p + q = 0$ and $pq = -1$. We could actually solve this equation, or just guess that $p = 1$ and $q = -1$ work. Indeed, it turns out that $x^2 - 1 = (x + 1)(x - 1)$.

Finally, $x^3 - x^2 - 2x$ is not a quadratic polynomial. However, the lack of a constant term means we can first factor out an $x$ term to get $x^3 - x^2 - 2x = x(x^2 - x - 2)$. We hence content ourselves to find $p, q$ such that $p + q = -1$ and $pq = -2$. This one is a bit tricky, but some thought reveals that $p = -2$ and $q = 1$ will do the trick, and indeed $(x - 2)(x + 1) = x^2 - x - 2$ so that

$$x^3 - x^2 - 2x = x(x - 2)(x + 1). \qquad \blacksquare$$

A useful factorization to keep in mind is the difference of $n$th powers formula:

$$x^n - a^n = (x - a)(x^{n-1} + x^{n-2}a + x^{n-3}a^2 + x^{n-4}a^3 + \cdots + a^{n-2}x + a^{n-1}),$$

where $n$ is a natural number.

Rational functions are quotients of polynomials; that is, they are functions which can be written as $f(x) = p(x)/q(x)$ where $p$ and $q$ are both polynomials. The following are examples of rational functions:

$$f(x) = \frac{x^2 + 2x + 1}{x - 1}, \qquad g(x) = \frac{1}{x^2 + 1}, \qquad h(x) = \frac{x^3 + 2x - 1}{4x^4 - x^2 + 13}.$$

## 1.4 Absolute Values

Absolute values are used to measure length and distance, which are naturally important. However, since lengths and distances should always be positive, the absolute value exhibits some subtlety that makes it difficult to work with. Here we'll give a brief review of the absolute value, and discuss how to manipulate it.

### 1.4.1 The Absolute Value

Given a number $x \in \mathbb{R}$ we would like to discuss its "distance" from the number 0. Naively, we would like to say something along the lines of "4 is the same distance from 0 as $-4$" or perhaps "2 is the same distance from 4 as $-1$ is from $-3$." Figure 1.9 illustrates this idea. The formal way to talk about the concept of length is with absolute values.

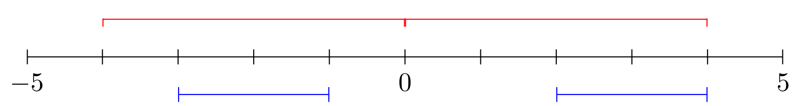



Figure 1.9: The real line from $-5$ to positive 5. We would like to define a system of measurement such that the red bars have the same length and the blue bars have the same length.

**Definition 1.13**

For $x \in \mathbb{R}$ we define the *absolute value* of $x$ as

$$|x| = \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases}.$$

Notice that the absolute value is always positive: If $x$ is already positive, the absolute value does not do anything, while if $x$ is negative we negate it again to make it positive. Geometrically, we may interpret $|x|$ as the distance from $x$ to 0. The distance from 4 to 0 is $|4| = 4$ while the distance from $-4$ to 0 is $|-4| = -(-4) = 4$. As we discussed above, this is precisely what we expected.

**Proposition 1.14: Properties of the Absolute Value**

If $a, b \in \mathbb{R}$ then

1. $|ab| = |a||b|$ (Multiplicative)
2. $|a+b| \leq |a| + |b|$ (Triangle Inequality)
3. $|a| = 0$ if and only if $a = 0$ (Non-degenerate).

### 1.4.2 Relation to Intervals

Instead of looking at the distance from $a$ to 0, we can look at the distance between $a$ and $b$, given by $|a-b|$. Therefore, we may use absolute values combined with inequalities to describe intervals. For example, consider the statement $|x-c| < a$. Using the definition of the absolute value, we can write this as

$$|x - c| = \begin{cases} x - c & x \geq c \\ c - x & x < c \end{cases}.$$

Now $|x-c| < a$ implies that both $x - c < a$ and $c - x < a$ for all values of $x$. If we multiply $c - x < a$ by $-1$ we get $x - c > -a$, which we may combine with $x - c < a$ to conclude that

$$|x - c| < a \quad \iff \quad -a < x - c < a.$$

We may read $|x-c| < a$ geometrically as

"The distance from $x$ to $c$ is less than $a$."

Intuitively, the set of all $x$ which satisfy this will lie in the interval $(c-a, c+a)$. We can show this more concretely by realizing that

$$\begin{aligned} |x - c| < a \quad &\iff \quad -a < x - c < a \\ &\iff \quad c - a < x < c + a \end{aligned} \tag{1.1}$$

**Example 1.15**

Find the intervals corresponding to all $x$ which satisfy the following inequalities:

$$|x| \leq 1, \qquad |2x - 5| < 3, \qquad |x + 7| > 5.$$

*Solution.* If $|x| \leq 1$ then $-1 \leq x \leq 1$ and this corresponds to the interval $[-1, 1]$. The next example is $|x-2| < 3$ and proceeding by the same argument in (1.1) we find that

$$\begin{aligned} |2x - 5| < 3 \quad &\iff \quad -3 < 2x - 5 < 3 \\ &\iff \quad 2 < 2x < 8 \\ &\iff \quad 1 < x < 4 \end{aligned}$$

so that the corresponding interval is $(1, 4)$.

Expressions of the form $|x+7| > 5$ will occur far less frequently than the examples considered above, but should still be solvable if we go back to the definition of absolute value. Intuitively we see that the $x$ which satisfy this will be a distance of at least 5 from $-7$; that is, $(-\infty, -12) \cup (-2, \infty)$. Let us check that this is the case.

The condition that $|x+7| > 5$ implies that both $x+7 > 5$ and $-x-7 > 5$. Solving the former for $x$ we find that $x > -2$ while the latter reveals that $x < -12$, precisely as we expected. ∎

### 1.4.3 Algebra with Inequalities

Working with absolute values within inequalities offers a new challenge. We use cases to remove the absolute value from our expression, then solve the corresponding inequality. However, we must consolidate our assumed case with our solution case to determine the true solution. This is best demonstrated through examples.

**Example 1.16**

Find all $x$ for which $|x+7| < 4x + 10$.

*Solution.* The equation $|x+7| < 4x+10$ is untenable in this form, because we cannot manipulate the absolute value directly. To deal with this, we examine the expression inside of the absolute value and determine where it changes from being positive to negative. Since $x + 7 = 0$ when $x = -7$, we break our problem into the case $x < -7$ where $|x+7| = -x-7$, and $x \geq -7$ where $|x+7| = x+7$.

**Case $x < -7$:** If we restrict ourselves to $x < -7$ then $|x+7| < 4x+10$ becomes

$$-x - 7 < 4x + 10.$$

Some quick manipulation shows us that $x > -17/5$, which combined with $x < -7$ tells us there are no solutions.

**Case $x \geq -7$:** In this case $|x+7| < 4x+10$ becomes $x+7 < 4x+10$. Some algebraic work shows that $x > -1$. Both $x > -1$ and $x \geq -7$ must be true at the same time, implying that $x > -1$ is the solution.

Combining the results from both cases, we see that $|x+7| < 4x+10$ if $x > -1$; that is, $x \in (-1, \infty)$. ∎

**Example 1.17**

Find all $x$ for which

$$|x-3| \geq |x+1| - 2. \tag{1.2}$$

*Solution.* The expression $|x-3|$ will change signs at $x = 3$ while $|x+1|$ will switch signs at $x = -1$. This implies that we should consider three cases: $x < -1$, $-1 < x < 3$, and $x > 3$.

**Case $x < -1$:** Equation (1.2) becomes

$$-x + 3 \geq -x - 1 - 2.$$

The $x$'s will cancel giving the expression $3 \geq -3$. This statement is always true, so $x < -1$ always satisfies the equation.

**Case $-1 < x < 3$:** In this case equation (1.2) becomes

$$-x + 3 \geq x + 1 - 2$$

which is solved to find $x \leq 2$. Hence $x$ must satisfy both $-1 < x < 3$ and $x \leq 2$ implying that $-1 < x \leq 2$.

**Case $x > 3$:** Now equation (1.2) becomes

$$x - 3 \geq x + 1 - 2$$

which yields $-3 \geq -1$, a false expression. This means that no $x$ in this region satisfies the equation.

Finally, we check the switch points $x = -1, 3$ themselves. Substituting $x = -1$ into (1.2) we get

$$|(-1) + 3| \geq |(-1) + 1| + 2 \quad \Rightarrow \quad 2 \geq 2$$

which is true, so that $-1$ satisfies the equation. On the other hand, $x = 3$ yields

$$|3 - 3| \geq |3 + 1| + 2 \quad \Rightarrow \quad 0 \geq 6$$

which is not true, so $x = 3$ does not satisfy the equation. Combining all of our information, the total solution is

$$\{x < -1\} \cup \{-1 < x \leq 2\} \cup \{-1\} = \{x \leq 2\}$$

or more concisely, the interval $(-\infty, 2]$. ■

## 1.5 Exponential Functions

As multiplication was motivated as a tactic for abbreviating $n$-fold sums, exponentiation was originally shorthand for $n$-fold products. That is, if $a$ is a real number and $n$ is a natural number, then we define $a^n$ as

$$\underbrace{a \times a \times \cdots \times a}_{n\text{-times}} = a^n. \tag{1.3}$$

As in the case of multiplication, we define the exponent for negative numbers; $a^{-n} = 1/a^n$. Exponentiation then satisfies the following rules: If $a, b$ are real numbers and $n, m$ are integers,

$$a^n a^m = a^{n+m}, \qquad (a^n)^m = a^{nm}, \qquad (ab)^n = a^n b^n.$$

### 1.5.1 Roots

To extend the idea of multiplication to rational numbers we exploited the notion of division, which is the "inverse" to multiplication. We will have to do something similar in order to exponentiate rational numbers.

**Definition 1.18**

If $a \geq 0$ is a real number and $n$ is a natural number, we define the *nth root* of $a$ to be the non-negative number $b$ such that $b^n = a$. When $n = 2$ we write $b = \sqrt{a}$, and when $n > 1$ we write $\sqrt[n]{a}$.

There are many subtleties in discussing roots, and in particular those subtleties can vary depending on the numbers we choose to plug into the definition. First, roots do represent a partial inverse to exponentiation. By definition, if $b = \sqrt[n]{a}$ then $b$ satisfies $b^n = a$; that is,

$$b^n = (\sqrt[n]{a})^n = a.$$

For this reason, we can write $\sqrt[n]{a} = a^{1/n}$. The properties of power laws then implies that

$$(a^{1/n})^n = a^{n/n} = a^1 = a.$$

The next is that the definition clearly states that $a$ must be a non-negative number. Why is this the case? Consider the instance in which we are asked to determine $b = \sqrt{-1}$, so that $b$ must satisfy $b^2 = -1$. We know that any number multiplied by itself must be non-negative, so there can be no solution to this equation.

Furthermore, if we consider the case $b = \sqrt{4}$ we see that there are two numbers satisfying $b^2 = 4$: $b = 2$ and $b = -2$. Since we would like roots to define functions we can only choose one $b$ as our solution, so we establish the convention of always choosing the positive solution.

The problems discussed above manifested when $n = 2$, and it turns out that these pathological examples only occur when $n$ is even. When $n$ is odd, there is no issue with taking $n$th roots of negative numbers, nor with the existence of multiple solutions. As an example, consider $b = \sqrt[3]{-8}$. There is a unique number, $b = -2$, such that $b^3 = -8$. We summarize our discussion below.

1. If $n$ is even and $a \geq 0$, then $b^n = a$ will have multiple solutions. To avoid ambiguity in defining $\sqrt[n]{a}$, we demand that $b$ must be non-negative.
2. If $n$ is even, it is impossible to define the $n^{th}$ root of a negative number.
3. If $n$ is odd, then neither 1 nor 2 apply; that is, there is a unique solution to $b^n = a$ for any $a \in \mathbb{R}$.

**Example 1.19**

Determine the values of $\sqrt{9}$ and $\sqrt[3]{-64}$.

*Solution.* Starting with $\sqrt{9}$, our goal is to find a positive integer $b$ such that $b^2 = 9$. We know that there will be multiple solutions since $n = 2$ is even, and indeed $b = 3$ and $b = -3$ both work. As our definition stipulates that $b$ must be non-negative, we take $b = 3$ and conclude that $\sqrt{9} = 3$.

On the other hand, as $n = 3$ is odd we know that $b = \sqrt[3]{-64}$ is the unique solution to $b^3 = -64$. A bit of trial and error shows that $(-4)^3 = -64$ and so $\sqrt[3]{-64} = -4$. ∎

By exploiting the identities given in Equation (1.3), we can immediately deduce the following for roots: If $a, b \in \mathbb{R}$ and $m, n \in \mathbb{Z}$ then

$$\sqrt[n]{a}\,\sqrt[m]{a} = \sqrt[mn]{a}^{\,m+n}\!\!\!, \qquad \sqrt[n]{\sqrt[m]{a}} = \sqrt[mn]{a}, \qquad \sqrt[n]{ab} = \sqrt[n]{a}\,\sqrt[n]{b}. \tag{1.4}$$

### 1.5.2 Logarithms

Given the equation $a^n = b$, we have discussed exponentiation and the process of taking roots. These ideas boil down to a two-out-of-three argument, so that if you are given two of variables solve for the third. For exponentiation, one is given $a$ and $n$ and told to determine $b$, while given $n$ and $b$ we may take $n$th roots to determine $a$. The remaining situation, given $a$ and $b$ determine $n$, is described by logarithms.

Why might we want to find such an $n$? There are many industrial reasons, the most often of which appear in pre-calculus courses as problems in finance. As an example, one is told that an asset appreciates at a fixed rate of 4% per annum and is tasked with determining the number of years until the asset's worth has doubled. This amounts to solving the equation $(1.04)^b = 2$, which we see is precisely the aforementioned problem which logarithms are designed to solve.

From a mathematical perspective, logarithms arise as the inverse to exponentiation. We saw that for a *fixed* natural number $n$ we could invert the process of exponentiation $x^n$ by taking an $n$th root. This is useful if we want to talk about inverses of polynomials. For example, if $f(x) = x^3$ then $f^{-1}(x) = \sqrt[3]{x}$. If we now fix the base and let the exponent vary, taking roots becomes untenable; in fact, our goal is to find the exponent itself! Logarithms are the solution to the inversion problem.

$a^n = b$

| | | |
|---|---|---|
| Exponentiation | Given:<br>Determine: | $a$ and $n$<br>$b$ |
| Roots | Given:<br>Determine: | $n$ and $b$<br>$a$ |
| Logarithms | Given:<br>Determine: | $a$ and $b$<br>$n$ |

Table 1: A description of the possible "two-out-of-three" situations arising from the equation $a^n = b$.

**Definition 1.20**

If $a$ and $b$ are positive numbers, we define $\log_a b$ (read as *the base-$a$ logarithm of $b$*) as the number $c$ satisfying $a^c = b$.

If you are unfamiliar with logarithms, the above definition can be a lot to take in. I would encourage you to take a second and parse Definition 1.20 until it starts to make sense.

**Example 1.21**

Compute $\log_2 32$, $\log_3 27$.

*Solution.* Let $c = \log_2 32$ in which case Definition 1.20 implies that $c$ must satisfy $2^c = 32$. You will hopefully recall that $2^5 = 32$, so $c = 5$ and we conclude that $\log_2 32 = 5$. Similarly, if $c = \log_3 27$ then $3^c = 27$. You can easily check that $3^3 = 27$ and so $\log_3 27 = 3$. ■

The manner in which we started this section should suggest that the logarithm is going to play the inverse role to exponentiation. Indeed, items 3 and 4 in the following proposition shed some light on the relationship between logarithms and exponentials.

**Proposition 1.22**

If $a$ and $b$ are positive real numbers with $a \neq 1$, then

1. $\log_a(1) = 0$
2. $\log_a(a) = 1$
3. $\log_a(a^b) = b$
4. $a^{\log_a(b)} = b$

These results are simple and you should make an attempt to prove the results on their own before looking at the proof. This will not only build confidence in working with logarithms, but also expand your comprehension of the subject.

*Proof.*

1. Set $c = \log_a(1)$ so that $a^c = 1$. Since $a \neq 1$ by hypothesis, it must be the case that $c = 0$. Thus $\log_a(1) = 0$ as required.
2. Similar to part 1, we know that $c = \log_a(a)$ satisfies $a^c = a$. It is not to hard to see that $c = 1$ is the only possible solution and hence $\log_a(a) = 1$.
3. Let $c = \log_a(a^b)$ so that $c$ satisfies $a^c = a^b$. It's not too hard to see that $c = b$ is the solution, so that $\log_a(a^b) = b$.
4. Let $c = \log_a(b)$ so that $a^c = b$. However, simply substituting our first expression of $c$ into the latter expression, we get $a^{\log_a(b)} = b$ as required. □

### 1.5.3 The Exponential and Logarithmic Functions

The procedure for extending exponentiation from $a^n$ for natural numbers $n$, to $a^x$ for real numbers $x$, is quite difficult. It requires that we either have access to the mathematics of sequences (which we will not cover), or integration (which is not covered until the second half of the course). As a result, you are going to have to take my word that such extensions exist.

We define an exponential function $f(x) = a^x$ whenever $a > 0$. This function has domain $\mathbb{R}$ and range $(0, \infty)$. There is a special value of the base $a$ known as *Euler's number*, denoted by $e$, with approximate value $e \approx 2.7182818284\ldots$. Unfortunately, the most intuitive definitions of this number require some notion of calculus, and so I only mention it here and define it later.

Figure 1.10 contains the graphs of several exponential functions. Notice that these satisfy the horizontal line test, and therefore should be invertible. The *logarithmic function with base* $a > 0$ is the function $g(x) = \log_a(x)$, which is designed to act as the inverse function for $f(x) = a^x$. Indeed, using items 3 and 4 of Proposition 1.23 we see that

$$a^{\log_a(x)} = x \quad \text{and} \quad \log_a(a^x) = x,$$

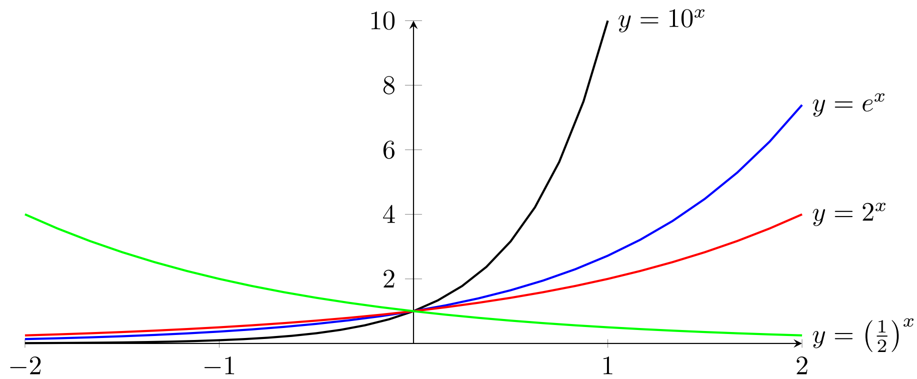



Figure 1.10: The graph of the exponential function for several choice of bases.

which is the relationship required of inverse functions. Recall that $a^x$ grows to infinity if $a > 0$ and shrinks to 0 if $0 < a < 1$. As such, we expect a similar dichotomy in the graphs of the logarithmic function. Since $a^x$ and $\log_a(x)$ are inverses, the graph of $\log_a(x)$ is just the reflection of $a^x$ about the line $y = x$ and is given in Figure 1.11. The domain of $\log_a(x)$ is $(0, \infty)$ while its range is all of $\mathbb{R}$.

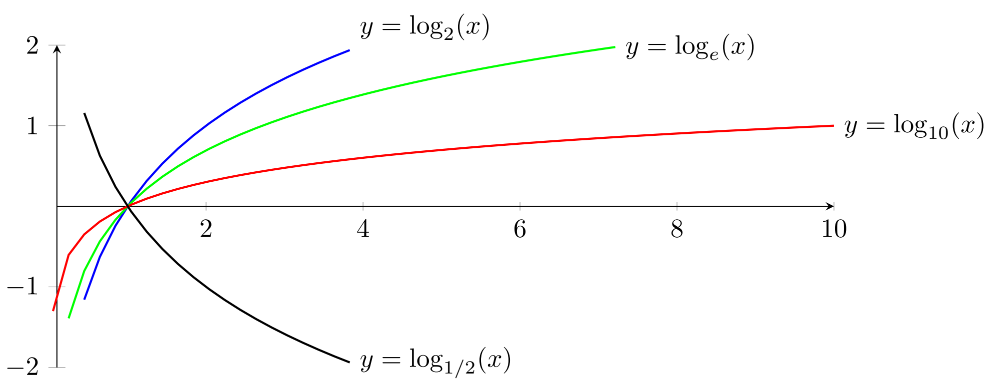



Figure 1.11: The graph of the logarithm function $\log_a(x)$ for different values of $a$, viewed as the reflection of the exponential function $a^x$ about the line $y = x$.

Given the close relationship between logarithms and exponents, it's not surprising that $e$ is a special base for the logarithmic function $\log_e(x)$. In fact, this function is so special that there are two competing mathematical conventions in writing it down. The first is to write $\ln(x)$, pronounced as "lawn of $x$," while the other is to simply omit the base and write $\log(x)$. The latter is typically used by mathematicians alone, while scientists and engineers prefer the $\ln(x)$ notation.

The next proposition gives a list of useful logarithmic identities, all of which may be proven by exploiting the relationship between the logarithm and the exponential. I will provide the proof of the most difficult result, but the rest are left as exercises for you.

**Proposition 1.23**

Let $d$ be any real number and $a$ be a positive number such that $a \neq 1$. For any $x, y > 0$ we then have

1. $\log_a(x^d) = d\log_a(x)$,
2. $\log_a(xy) = \log_a x + \log_a y$
3. $\log_a(x/y) = \log_a x - \log_a y$,
4. $\log_a b = \dfrac{\log_d b}{\log_d a}$.

*Proof.* The proofs of 1, 2, and 3 are exercises in applying the appropriate exponential identity and are left to you. I will prove 4 here. Define

$$\begin{aligned} c &= \log_a b & c_1 &= \log_d b & c_2 &= \log_d a \\ a^c &= b & d^{c_1} &= b & d^{c_2} &= a \end{aligned} \tag{1.5}$$

Starting with $a^c = b$, we substitute the latter two expressions in (1.5) to get

$$\begin{aligned} a^c &= b \\ (d^{c_2})^c &= (d^{c_1}) && \text{since } a = d^{c_2} \text{ and } b = d^{c_1} \\ d^{c_2 \times c} &= d^{c_1}, \end{aligned}$$

which implies that $c_2 \times c = c_1$. Solving for $c$ we get $c = c_1/c_2$ or rather

$$\log_a b = \frac{\log_d b}{\log_d a}$$

as required. □

## 1.6 Sigma Notation and Geometric Series

Sigma notation is used to make complicated sums easier to write down. In particular, we use a summation index to iterate through elements of a list and then sum them together. Consider the expression

$$\boxed{\sum_{i=n}^{m} r_i} \tag{1.6}$$

which is read as "the sum from $i = n$ to $m$ of $r_i$." The element $i$ is known as the *dummy* or *summation* index, $n$ and $m$ are known as the *summation bounds*, and $r_i$ is the *summand*. In order to decipher this cryptic notation, we adhere to the following algorithm:

1. Set $i = n$ and write down $r_i$;

2. Add 1 to the index $i$ and add $r_i$ to the current sum;
3. If $i$ is equal to $m$ then stop, otherwise go to step 2 and repeat.

For those computer savvy students out there, this is nothing more than a for-loop. Interpreting (1.6) we thus have

$$\sum_{i=n}^{m} r_i = r_n + r_{n+1} + r_{n+2} + \cdots r_m.$$

**Example 1.24**

Set $r_1 = 5, r_2 = -8, r_3 = 4$. Compute $\displaystyle\sum_{i=1}^{3} r_i$.

*Solution.* Via our discussion above, we may write the summation explicitly as

$$\sum_{i=1}^{3} r_i = r_1 + r_2 + r_3 = 5 + (-8) + 4 = 1. \qquad \blacksquare$$

The $r_i$ could be a collection of unrelated numbers as in Example 1.24, but they could be a "function" of the index variable as follows:

**Example 1.25**

Compute $\displaystyle\sum_{i=1}^{4} (2i+1)$.

*Solution.* Following our algorithm, we start by setting $i = 1$ and then evaluating the summand. I will write out the steps in slightly more detail than usual to illustrate the process:

$$\begin{aligned}\sum_{i=1}^{4} (2i+1) &= (2i+1)_{i=1} + (2i+1)_{i=2} + (2i+1)_{i=3} + (2i+1)_{i=4} \\ &= (2\cdot 1+1) + (2\cdot 2+1) + (2\cdot 3+1) + (2\cdot 4+1) \\ &= 3+5+7+9 \\ &= 24. \qquad \blacksquare\end{aligned}$$

Sometimes we can find closed form expressions for summations. You will not be expected to memorize the following, but they are nonetheless important identities:

$$\begin{aligned} &\sum_{j=1}^{n} 1 = n & &\sum_{j=1}^{n} j = \frac{n(n+1)}{2} \\ &\sum_{j=1}^{n} j^2 = \frac{n(n+1)(2n+1)}{6} & &\sum_{j=1}^{n} j^3 = \left[\frac{n(n+1)}{2}\right]^2 \end{aligned} \tag{1.7}$$

Additionally, summations are linear, in that

$$\left[\sum_{i=n}^{m} a_i\right] + \left[\sum_{i=n}^{m} b_i\right] = \sum_{i=n}^{m}(a_i + b_i), \qquad c\left[\sum_{i=n}^{m} a_i\right] = \sum_{i=n}^{m}(ca_i).$$

Note that the upper and lower bounds of the summation are the same in every summation.

Using linearity and the identities in (1.7) we can redo Example 1.25 with a general upper bound, to find

$$\begin{aligned}\sum_{i=1}^{n}(2i+1) &= 2\left(\sum_{i=1}^{n} i\right) + \left(\sum_{i=1}^{n} 1\right)\\ &= 2\frac{n(n+1)}{2} + n\\ &= n(n+2).\end{aligned}$$

Plugging in $n = 4$ we get 24, just as we found in Example 1.25.

**Remark 1.26** For any positive integer $p$, there is a closed form expression for

$$\sum_{i=1}^{n} i^p$$

but these expressions become more difficult as $p$ becomes larger. Luckily, there is a standard way of deriving the closed form for any $p$ using the *Bernoulli polynomials*, which are popular objects in the study of number theory but are tricky to define.

There are other summations which also admit closed form expressions, which are not evaluated as easily as the examples above.

**Example 1.27**

Guess a closed form expression for the summation $\displaystyle\sum_{i=1}^{n} \frac{1}{i^2+i}$.

*Solution.* We will try a few values of $n$, such as $n = 1, 2, 3, 4$; to see if we can spot a pattern. Indeed,

$$\begin{aligned}n = 1: &\qquad \sum_{i=1}^{1} \frac{1}{i^2+i} = \frac{1}{2}\\ n = 2: &\qquad \sum_{i=1}^{2} \frac{1}{i^2+i} = \frac{1}{2} + \frac{1}{6} = \frac{2}{3}\\ n = 3: &\qquad \sum_{i=1}^{3} \frac{1}{i^2+i} = \frac{2}{3} + \frac{1}{12} = \frac{3}{4}\\ n = 4: &\qquad \sum_{i=1}^{4} \frac{1}{i^2+i} = \frac{3}{4} + \frac{1}{20} = \frac{4}{5}.\end{aligned}$$

Were we to guess, it looks as though

$$\sum_{i=1}^{n} \frac{1}{i^2+i} = \frac{n}{n+1}, \tag{1.8}$$

and indeed this is true. Can you prove it? ■

### 1.6.1 Geometric Series

A geometric series is a series where every term of the series is a multiple of the previous one. For example,

$$a_0 = 1,\ a_1 = \frac{1}{2},\ a_2 = \frac{1}{4},\ a_3 = \frac{1}{8},\ a_4 = \frac{1}{16}, \ldots$$

satisfies the relation $a_n = \frac{1}{2}a_{n-1}$. We can write such series as

$$\sum_{k=0}^{\infty} ar^k,$$

where $a = a_1$ and $r = a_n/a_{n-1}$ for any $n$.

**Theorem 1.28**

If $r$ is a real number and $n$ is a positive integer, then

$$\sum_{k=0}^{n} r^k = \frac{1-r^{n+1}}{1-r} \quad \text{and} \quad \sum_{k=1}^{n} r^k = \frac{r(1-r^n)}{1-r}.$$

*Proof.* For brevity, let $S_n = 1 + r + r^2 + \cdots + r^n$ denote the sum. Multiplying by $1 - r$ we get

$$\begin{aligned}(1-r)S_n &= 1 + r + r^2 + r^3 + r^4 + \cdots + r^n \\ &\qquad - r - r^2 - r^3 - r^4 + \cdots - r^n - r^{n+1} \\ &= 1 - r^{n+1}.\end{aligned}$$

Solving for $S_n$ gives $S_n = \dfrac{1-r^{n+1}}{1-r}$. □

**Example 1.29**

Determine the sum of the series $\displaystyle\sum_{k=0}^{10} 2^k$ and $\displaystyle\sum_{k=1}^{8} \pi^{2k}$.

*Solution.* The common ratio for the first series is $r = 2$. Substituting into our formula gives

$$\sum_{k=0}^{10} 2^k = \frac{1-2^{11}}{1-2} = 2047.$$

The second series looks like

$$\sum_{k=1}^{8} \pi^{2k} = \pi^2 + \pi^4 + \pi^6 + \cdots + \pi^{16}$$

so the common ratio is $r = \pi^2$. Evaluating our formula produces

$$\sum_{k=1}^{8} \pi^{2k} = \frac{\pi^2(1-\pi^{16})}{1-\pi^2}. \qquad \blacksquare$$

**Theorem 1.30**

For any $|r| < 1$ we have

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r} \tag{1.9}$$

and the series diverges otherwise.

*Proof.* Since $|r| < 1$, we have that $r^{n+1} \xrightarrow{n\to\infty} 0$, so

$$\sum_{k=0}^{\infty} r^k = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{1-r^{n-1}}{1-r} = \frac{1}{1-r}. \qquad \square$$

An interesting if non-rigorous argument that proves the same result is to see that

$$\begin{aligned}(1-r)(1+r+r^2+r^3+\cdots) &= 1 + r + r^2 + r^3 + \cdots \\ &\qquad - r - r^2 - r^3 - \cdots \\ &= 1.\end{aligned}$$

In an entirely formal sense (that is, treating $r$ purely as a symbol without assigning it any value) we see that $(1-r)$ is the multiplicative inverse of $(1+r+r^2+\cdots)$, giving the desired results as well.

**Example 1.31**

Determine the limit $\displaystyle\sum_{k=2}^{\infty} \frac{1}{2^k}$.

*Solution.* Notice that our summation index begins at 2, and not 0 as in (1.9). We can fix this by realizing that

$$\sum_{k=2}^{\infty} \frac{1}{2^k} = \sum_{k=0}^{\infty} \frac{1}{2^k} - 1 - \frac{1}{2} = \frac{1}{1-1/2} - 1 - \frac{1}{2} = 2 - 1 - \frac{1}{2} = \frac{1}{2}. \qquad \blacksquare$$

**Example 1.32**

Determine the value of the series $\sum_{n=0}^{\infty} \frac{(-2)^n}{e^n}$.

*Solution.* Here our ratio is $r = -2/e$. Since $e \approx 2.7182$ we know $|r| < 1$. By (1.9) we have

$$\sum_{n=0}^{\infty} \frac{(-2)^n}{e^n} = \frac{1}{1-(-2/e)} = \frac{e}{e+2}. \qquad \blacksquare$$

**Example 1.33**

Determine the value of the series $\sum_{k=10}^{\infty} \frac{-3}{4^k}$.

*Solution.* The $-3$ can be pulled outside of the sum. As our index does not start at 0 we could subtract the first 10 terms, but this is rather onerous. Instead, we can write

$$\begin{aligned}\sum_{k=10}^{\infty} \frac{1}{4^k} &= \frac{1}{4^{10}} + \frac{1}{4^{11}} + \frac{1}{4^{12}} + \frac{1}{4^{13}} + \cdots \\ &= \frac{1}{4^{10}}\left(1 + \frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + \cdots\right) \\ &= \frac{1}{4^{10}} \sum_{k=0}^{\infty} \frac{1}{4^k}.\end{aligned}$$

Putting this all together we get

$$\sum_{k=10}^{\infty} \frac{-3}{4^k} = \frac{-3}{4^{10}} \sum_{k=0}^{\infty} \frac{1}{4^k} = \frac{-3}{4^{10}} \frac{1}{1-1/4} = -\frac{1}{4^9}. \qquad \blacksquare$$

## 1.7 Exercises

1-1. Define the sets

$$A = \mathbb{Z}, \quad B = \{0\} \cup (1,2), \quad C = (-\infty, 0) \cup (0, \infty), \quad D = \left\{x \in \mathbb{R} : x^2 < 2 \text{ or } x < 0\right\}.$$

Compute each of the following sets:

(a) $A \cup B$
(b) $A \cup C$
(c) $A \setminus B$
(d) $B \cap C$
(e) $C \cap D$
(f) $C \setminus B$
(g) $B \cap D$
(h) $(C \cup B) \cap A$
(i) $D \setminus C$
(j) $(A \cup B) \cap (C \cup D)$
(k) $(A \cap C) \cup (C \cap D)$
(l) $(A \cup D) \setminus C$

1-2. If $f$ is a function, a fixed point of $f$ is any point $x_0$ which satisfies $f(x_0) = x_0$. Let $f(x) = -x^5 - 10x^3 + e^x$ and let $A$ be the set of real numbers which are fixed points for $f$. Write down $A$ in terms of set builder notation. *Note:* You do not need to find the fixed points, just write down a valid set-builder formula.

1-3. Define the sets

$$A = \{-1, 0, 1\}, \quad B = \{1, 2, 3, \ldots, \} = \mathbb{N}, \quad C = [0, 1], \quad \text{and} \quad D = \emptyset.$$

Compute each of the following sets, and simplify your answer as much as possible.

(a) $A \cap C$
(b) $A \cap B$
(c) $C \cap D$
(d) $A \setminus C$
(e) $C \setminus A$
(f) $(A \cap B) \setminus C$
(g) $(A \cup C) \setminus B$

1-4. Define the set $A = \{x \in \mathbb{Z} : (x - 2)(x + 1) \leq 0\}$. What is $A$?

1-5. Let $A = \{x \in \mathbb{R} : x \geq -10\}$ and $B = (-\infty, 2) \cup (5, \infty)$. Write $A$ is an interval or union of intervals.

1-6. We define the sets $A$, $B$, and $C$ as follows:

$A = \{x \in \mathbb{R} : e^x < 100\}$,
$B =$ the values of $x$ such that the function $f(x) = x^2 - 4$ is non-negative,
$C =$ the set of natural numbers which, when squared, are strictly greater than 10.

(a) Write the sets $B$ and $C$ using set-builder notation. In particular, do not use any intervals to write out $B$ or $C$.
(b) Write the sets $A$ and $B$ as intervals or unions of intervals as appropriate.
(c) Write the sets $A \cap B$ and $A \setminus C$ as intervals or unions of intervals as appropriate. Simplify your answer as much as possible by expressing your answer using as few intervals as possible.

1-7. (a) Define the sets

$$A = \left\{x \in \mathbb{R} : x^2 > 100\right\}, \quad B = \{2^n : n \in \mathbb{N}\} = \{1, 2, 4, 8, 16, \ldots\}, \quad C = \mathbb{Q} \cap (0, 1),$$

all thought of as subsets of $\mathbb{R}$. Determine the sets $A \cup B$, $A \cup C$, $B \setminus A$, and $A \setminus C^c$. You may use unions, intersections, and set-difference in your answers, but your solution should be simplified as much as possible.

(b) Indicate whether the statement is true or false. You do not need to justify your answer:
i. If $A \subseteq B$, then $A \cup B = B$.
ii. If $A, B$ are subsets of $\mathbb{R}$ and $A \cup B = \mathbb{R}$, then either $A = \mathbb{R}$ or $B = \mathbb{R}$.
iii. If $A \subseteq B^c$ then $A \cap B = \emptyset$.

1-8. Determine the domain of each given function.

(a) $f(x) = \sqrt{25 - x^2}$

(b) $g(x) = \dfrac{1}{e^{x+2} - 1}$

(c) $h(x) = \ln(|x| - 2)$

(d) $\xi(x) = \dfrac{1}{x^2 - x - 12}$

1-9. Determine $f \circ g$ for each given $f$ and $g$.

(a) $f(x) = 2x - 3, g(x) = x^2 + 1$

(b) $f(x) = xe^x, g(x) = e^x$

(c) $f(x) = e^x, g(x) = x\ln(x^2 + 1)$

(d) $f(x) = \dfrac{x}{x^2 + 1}, g(x) = |x - 1|$

1-10. Determine the inverse of the following functions:

(a) $f(x) = 4x - 3$

(b) $f(x) = 1/x$

(c) $f(x) = \dfrac{x - 2}{1 - 3x}$

(d) $f(x) = \dfrac{2x - 1}{4 + 3x}$

1-11. Find all $x$ which satisfy the following identities

(a) $|3x - 5| = 14$

(b) $|2x + 5| \geq 7$

(c) $|4x + 32| > -1$

(d) $|3x - 4| = |2x + 5|$

(e) $2|x - 3| - 3|x - 2| < 1$

1-12. Given each condition on $f$ and $g$, determine if

i. $f + g$

ii. $fg$

iii. $f \circ g$

is even, odd, or neither.

(a) $f$ is even, $g$ is even

(b) $f$ is even, $g$ is odd

(c) $f$ is odd, $g$ is even

(d) $f$ is odd, $g$ is odd

1-13. For general values of $a, b, c, d$, determine the inverse of $f(x) = \dfrac{ax + b}{cx + d}$. What condition on $a, b, c, d$ is necessary to ensure that the function is invertible?

1-14. Simplify the following expressions as much as possible. Where numbers are involved, do not use a calculator.

(a) $\log_8(4)$

(b) $3^2 + 3^2 + 3^2$

(c) $9^{\log_3(x^2)}$.

(d) $\ln(e^x e^y)$ for real numbers $x, y$.

(e) $6^{10} 2^{-8} 3^{-9}$

(f) $e^{x\ln(x) + (x-1)\ln(x)}$.

(g) $\log_{a^b}(a^c)$ for positive real numbers $a, b, c$.

(h) $\ln\left(\dfrac{1}{x}\sqrt{\dfrac{x}{4}}\right)$.

1-15. Solve each equation for $x$.

(a) $e^{\ln(x^2)} = 16$,
(b) $e^{x^2-x-12} = 1$,
(c) $\log(x^3 - 3) = 1$,
(d) $\log_2(x) + \log_2(x^2) = 8$,
(e) $\log_{10}(x-3) + \log_{10}(x-5) = 1$
(f) $\ln(x+5) = \ln(3x+3) + 1$
(g) $3 - x = 2\log_2(5) - \log_2(3^x - 2^{x-4})$

1-16. Write the following series in sigma notation:

(a) $4 + 6 + 8 + 10 + 12 + \cdots + 22 + 24$
(b) $3 + 7 + 11 + 15 + \cdots + 27$
(c) $1 + \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \frac{1}{81}$
(d) $3 - 6 + 12 - 24 + 48 - 192$
(e) $\frac{1}{2\sqrt{2}} + \frac{1}{8} + \frac{1}{16\sqrt{2}} + \frac{1}{64}$

1-17. Solve the equation $A = R\left[\frac{1-(1+r)^{-n}}{r}\right]$ for $n$.

1-18. If $a, r$ are two real numbers, we know that $\sum_{k=0}^{n} ar^k = \frac{a(1-r^{n+1})}{1-r}$. Use this to determine the following sums:

(a) $\sum_{k=1}^{n} ar^k$
(b) $\sum_{k=0}^{n-1} ar^k$
(c) $\sum_{k=1}^{n-1} ar^k$
(d) $\sum_{k=0}^{n} ar^{2k}$

1-19. Evaluate each of the following sums:

(a) $\sum_{k=0}^{10} (2^k + 3^{k+2})$
(b) $\sum_{k=1}^{4} (-3)^{3k}$
(c) $\sum_{k=-2}^{5} \left[4^{k+3} + 4\right]$

1-20. (a) Suppose $r$ is a real number with $|r| < 1$. Show that

$$\sum_{k=1}^{\infty} r^k = \frac{r}{1-r}, \quad \sum_{k=2}^{\infty} r^k = \frac{r^2}{1-r}, \quad \sum_{k=3}^{\infty} r^k = \frac{r^3}{1-r}.$$

(b) Hypothesize a formula for $\sum_{k=m}^{\infty} r^k$ and show that your answer is correct.

1-21. Let $A = \sum_{\ell=4}^{8} 10^\ell$ and $B = \sum_{j=1}^{5} 7^j$. Write $A + B$ using a single summation.

1-22. Suppose you know the following three facts:

$$\sum_{r=1}^{20} a_r = 52, \quad \sum_{r=1}^{11} a_r = 23, \quad \sum_{r=11}^{20} a_r = 37.$$

What is the value of $a_{11}$?

1-23. Is the following statement true or false? If it's true, show that it's true. If it's false, give a counterexample.

$$\sum_{k=1}^{n} [a_n b_n] = \left[\sum_{k=1}^{n} a_n\right] \left[\sum_{k=1}^{n} b_n\right].$$

# 2 Financial Mathematics

In this section we'll take a look at the *time value of money*, and the mathematics used to compare money at different times. The mathematics itself only relies on high school mathematics, so our focus instead is on the concepts of "moving money through time."

## 2.1 Compounding Interest

Given a *principal* (initial investment) $P$ and an interest rate $r$, compounding interest is the notion that interest payments themselves can accumulate further interest. As a simple example, suppose that \$100 is invested in an account which yields 8% on the account balance at the end of each year. Here $P = 100$ and $r = 0.08$. After one year, the amount in the account is

$$\underbrace{100}_{\text{Principal}} + \underbrace{(100 \times 0.08)}_{\text{Interest}} = 100(1.08) = \$108.00. \tag{2.1}$$

Allowing this money to sit for another year – without supplementing the principal – the yield after year two is

$$108 + \underbrace{(108 \times 0.08)}_{\text{Interest}} = 108(1.08) = \$116.64. \tag{2.2}$$

Note the difference in growth between the two years: In the first year the interest contributed \$8 to the account, but in year two the account increased \$8.64. This is the notion of *compounding interest*: The interest in year two accumulates on both the principal ($100 \times 0.08 = \$8$), *and* the interest for year one ($8 \times 0.08 = \$0.64$). Similarly, year three will see the \$8 increase from the principal, and an additional $\$1.33 = 16.64 \times 1.08$ arising from interest generated after year one and two. Below is a chart showing the growth over a 10 year period.

| Year | Initial | Final | Difference |
|---|---|---|---|
| 1 | \$100.00 | \$108.00 | \$8.00 |
| 2 | \$108.00 | \$116.64 | \$8.64 |
| 3 | \$116.64 | \$125.97 | \$9.33 |
| 4 | \$125.97 | \$136.05 | \$10.08 |
| 5 | \$136.05 | \$146.93 | \$10.88 |
| 6 | \$146.93 | \$158.69 | \$11.75 |
| 7 | \$158.69 | \$171.38 | \$12.69 |
| 8 | \$171.38 | \$185.09 | \$13.71 |
| 9 | \$185.09 | \$199.90 | \$14.81 |
| 10 | \$199.90 | \$215.89 | \$15.99 |

From this chart, we see that \$100 has a 10 year *future value* of \$215.89. Financial and economic decisions need to take these future values into consideration, since making a decision now which costs you \$100 has an opportunity cost.

### 2.1.1 The Mathematics of Simple Compounding Interest

**Simple Compounding Interest I:** If a principal $P$ is compounded at an interest rate $r$, the value of the investment $S$ after one compounding period is

$$S = P(1+r). \tag{2.3}$$

If you don't see where this equation comes from, revisit Equations 2.1 and 2.2. To determine the value of the account after one compounding period, we compute the interest $Pr$ and add it to the original principal:

$$S = \underbrace{P}_{\text{Principal}} + \underbrace{Pr}_{\text{interest}} = P(1+r).$$

**Example 2.1**

Suppose a principal $P$ receives compounding interest at the end of each month at an interest rate $r$. Determine the value of the account $S_N$ after $N$ months.

*Solution.* The question right now is rather abstract, so let's break the problem into something more digestible. Let's try to determine the value of the account after 2 months, then after 3 months, and see if there is a pattern.

From Equation (2.3) we know that after one month, the value of the account is

$$S_1 = P(1+r).$$

To determine the value of the account after two months, there are two ways we can proceed. The first is just to compute $S_2$ from scratch, similar to Equation (2.2). Computing the new interest and adding it to the previous balance we get

$$S_2 = \underbrace{P(1+r)}_{\substack{\text{Previous}\\ \text{balance}}} + \underbrace{P(1+r)r}_{\text{Interest}} = P(1+r)(1+r) = P(1+r)^2.$$

This was a lot of work, and will not scale well when we have to compute many compounding periods.

The second technique is to change how we think about the problem. The value in the account at the end of the first compounding period is $S_1 = P(1+r)$, and now we wish to compound again. Mathematically, this is equivalent to saying that our original investment is $P(1+r)$ dollars. Now by compounding at a rate $r$ and applying Equation (2.3) we get

$$S_2 = \underbrace{S_1}_{\substack{\text{New}\\ \text{Principal}}}(1+r) = \underbrace{P(1+r)}_{S_1=P(1+r)}(1+r) = P(1+r)^2.$$

This second technique is mathematically faster, though it requires a better conceptual understanding of the problem.

We won't use the first technique ever again, so let's get comfortable using the second strategy. To determine the value in the account after a third compounding period, we think of $S_2$ as our new principal, and apply Equation 2.3:

$$S_3 = S_2(1+r) = P(1+r)^2(1+r) = P(1+r)^3.$$

Let's summarize our results below:

$$S_1 = P(1+r)^1, \quad S_2 = P(1+r)^2, \quad S_3 = P(1+r)^3.$$

If we continue this process, each time we compound we multiply by $(1+r)$, thus $S_N = P(1+r)^N$. ■

The fact that we compounded monthly in Example 2.1 made no difference to the answer. If we had compounded daily and had asked for the value of the account after $N$ days, we would get the same answer. Similarly, if we had compounded yearly, the value of the account after $N$ years would be the same.

**Simple Compounding Interest II:** If a principal $P$ is compounded at an interest rate $r$, the value of the investment $S$ after $N$ compounding periods is

$$S = P\underbrace{(1+r)(1+r)\cdots(1+r)}_{N\text{-times}} = P(1+r)^N. \tag{2.4}$$

Let's use this to double check the math from our motivating problem at the beginning of the chapter. Suppose $P = \$100$ and $r = 0.08$. After $N = 10$ compounding periods, the future value of our account should be

$$S_{10} = P(1+r)^N = 100(1+0.08)^{10} = 215.89$$

which is precisely the answer we found earlier. Notice as well that we didn't have to compute $S_1, S_2, \ldots, S_9$ in order to compute $S_{10}$, and this formula allowed us to get directly to the answer.

Don't fall into the trap of relying on formulas! While Equation 2.4 is useful and will be used often, there are times when it's necessary to return to basics.

**Example 2.2**

Suppose a principal $P$ is invested on June 1st. At the end of each odd day, the account balance is compounded at an interest rate $r$. At the end of each even day, the account balance is compounded at an interest rate $s$. What is the value of the account at the end of June? (June has 30 days.)

*Solution.* Once we determine a pattern, we'll be able to solve this question quickly. Let's start by applying what we know. At the end of the first day we compound at a rate of $r$, so after day one we have $S_1 = P(1+r)$ dollars.

Now $S_1$ is the balance in the account at the start of day two, wherein we now compound at a rate of $s$, so

$$S_2 = \underbrace{S_1}_{\substack{\text{New}\\ \text{balance}}} (1+s) = P(1+r)(1+s).$$

At the end of the third day, $S_2$ is the balance and we compound at a rate $r$, so

$$S_3 = S_2(1+r) = [P(1+r)(1+s)]\,(1+r) = P(1+r)^2(1+s).$$

We'll do one more day just to be certain. At the end of day four, the account has $S_3$ dollars and we compound at a rate $s$, so

$$S_4 = S_3(1+s) = \left[P(1+r)^2(1+s)\right](1+s) = P(1+r)^2(1+s)^2.$$

So we seem to pick up a $(1+r)$ for every odd day that has passed, and a $(1+s)$ for every even day that has passed. At the end of June, there were exactly 15 even days and 15 odd days, so $S_{30} = P(1+s)^{15}(1+r)^{15}$. ■

A helpful visualization technique is to plot a timeline of the compounding periods. In the case of Example 2.1, our terms were monthly and each month incurred an interest of $r$, so each time the principal $P$ is moved through a compounding period, it picks up a factor of $(1+r)$ as seen in Figure 2.1.

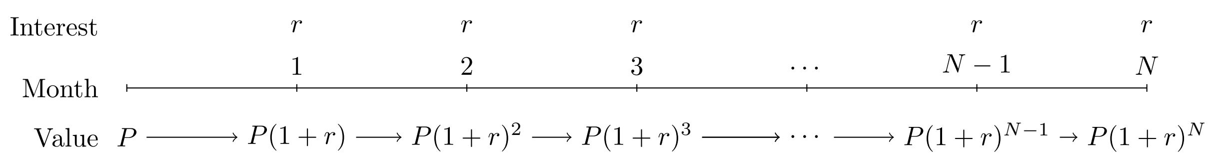



Figure 2.1: Moving a principal to the future through compounding periods incurs a multiplicative factor of $1+r$.

Precisely the same diagram helps us with Example 2.2, only this time the interest rates changes. See Figure 2.2.

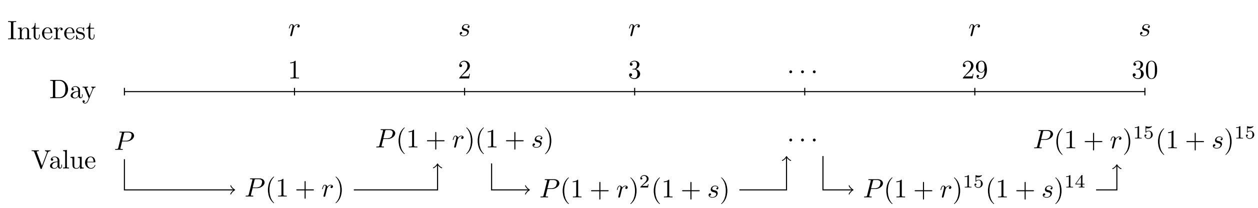



Figure 2.2: Even with altnerating interest rates, this visualization scheme still works.

### 2.1.2 Annual Percentage Rates

In reality, things like loans are compounded multiple times per year. In this case, one is often quoted an *Annual Percentage Rate* (APR) $r$ and told how often the principal compounds in a year.

The per-term interest rate is the $r$ divided evenly over each of the terms. For example, if you are quoted an APR of $r = 0.024$ compounding monthly, you take that $r = 0.024$ and divide it into 12 equal pieces – one for each month of the year. Thus each month you compound at a rate of $0.024/12 = 0.002$.

From another perspective, when computing the return on an investment one might be quoted the *nominal rate*, which is the amount by which the investment grows in a year. Computationally, the nominal rate and APR are equivalent. Future values are often affected by inflation, and subtracting the effect of inflation from the nominal rate results in the *real rate of return*. We won't discuss this much during the course.

**Example 2.3**

Suppose a principle $P$ is quoted as compounding at an Annual Percentage Rate $r$, compounded $t$-times per year. Determine the future value of the principal after $n$ years.

*Solution.* When quoted an APR of $r$ compounding $t$-times per year, Equation (2.4) still holds but requires a slight modification. The interest used for each period is $r/t$ – since we divide $r$ evenly into the number of compounding periods – and over $n$ years there are $N = nt$ compounding periods. This gives us the following formula:

$$S = P\left(1 + \frac{r}{t}\right)^{tn}. \qquad \blacksquare$$

The result of Example 2.3 is useful enough that it warrants special attention. Again, be careful not to just memorize this formula, and that you understand from where it comes!

**Compounding Interest:** Suppose $P$ is a principal, prescribed to grow at an annual percentage rate $r$, compounded $t$ times per year. The value of the investment $S$ after $n$ years is

$$S = P\left(1 + \frac{r}{t}\right)^{tn}. \tag{2.5}$$

Common compounding terms include yearly ($t = 1$), semi-annually ($t = 2$), quarterly ($t = 4$), and monthly ($t = 12$). There are institutions which compound daily ($t = 365$), but these are rare.

**Example 2.4**

An equity investment of \$10,000 grows at an annual percentage rate of 6%, compounded monthly. Determine the value of the investment after 6 months, 1 year, and 3 years.

*Solution.* Right away we know $P = 10000$, $t = 12$, and $r = 0.06$. Now we can directly apply the formula, or we can compute things from scratch. The monthly interest rate is $r/t = 0.06/12 = 0.005$. Thus after 6 months we get

$$S = 10000\,(1 + 0.005)^6 = \$10303.78.$$

Note that we were able to do this without referring to a formula at all. However, if you'd like to use Equation (2.5), we need to recognize that 6 months is half a year, so $n = 1/2$:

$$S = 10000\left(1 + \frac{0.06}{12}\right)^{12\times\frac{1}{2}} = 10000\,(1 + 0.005)^6 = 10303.78.$$

For $n = 1$ and $n = 2$ we thus get

$$10000\left(1 + \frac{0.06}{12}\right)^{12\times 1} = 10000\,(1 + 0.005)^6 = 10616.78$$

and

$$10000\left(1 + \frac{0.06}{12}\right)^{12\times 3} = 10000\,(1 + 0.005)^{36} = 11966.81. \qquad \blacksquare$$

Look at the 1-year answer to Example 2.4, which we found to be \$10616.78. A quick computation tells us that if we had compounded \$10000 at 6% compounding yearly, then after a single year we would have $10000(1 + 0.06) = 10600$ dollars. This is a different answer than if we had compounded monthly. The 1-year value of \$10000 compounded in different intervals is given in Table 2 below:

| $t$ | 1 | 2 | 3 | 4 | 6 | 12 |
|---|---|---|---|---|---|---|
| $S$ | \$10600.00 | \$10609.00 | \$10612.08 | \$10613.64 | \$10615.20 | \$10616.78 |
| $r_a$ | 6% | 6.09% | 6.1208% | 6.1364% | 6.1520% | 6.1678% |

Table 2: The future value of a principal $P = 10000$ invested at an APR of $r = 0.06$ for 1 year, over different compounding periods $t$. The value $r_a$ is the effective annual rate.

Of note is that the value of $S$ increases as $t$ increases; that is, *the effect of compounding interest is greater the more times the principal is compounded.* The annual percentage rate can thus be misleading in terms of the actual rate. For this reason, we define the *effective annual rate of return* as the rate which yields an equivalent return when compounded only once ($t = 1$). In the above example, if we compound \$10000 monthly at 6%, we make \$10616.79. The effective annual rate is the rate $r_a$ such that if we had *only compounded once*, we would have returned the same amount of money:

$$10000(1 + r_a) = 10616.79 \quad \Rightarrow \quad r_a = 0.061679.$$

**Example 2.5**

Suppose a principal $P$ is invested at an Annual Percentage Rate of $r$ compounded $t$ times annually. Determine the effective annual rate.

*Solution.* Let $r_a$ be the effective annual rate. According to Equation (2.5), the value of our account after one year is

$$S = P\left(1 + \frac{r}{t}\right)^t.$$

The question is to therefore determine the value $r_a$ so that $P(1+r_a) = S$. Setting these two equations equal to one another and solving for $r_a$ we get

$$P(1+r_a) = P\left(1+\frac{r}{t}\right)^t \quad \Rightarrow \quad r_a = \left(1+\frac{r}{t}\right)^t - 1. \tag{2.6}$$

Note that this does not depend on the initial principal $P$. ■

The benefit of the Equation (2.6) is that we don't need to first compute the future value of the account, and instead we can compute $r_a$ directly:

$$r_a = \left(1+\frac{0.06}{12}\right)^{12} - 1 \approx 0.061678,$$

showing that compounding monthly yields an effective rate of 6.1678%. We can even check that our answer is correct, since

$$\underbrace{10000(1+r_a)^1 = 10616.78}_{\text{compounded once at } r_a}, \qquad \underbrace{10000\left(1+\frac{r}{12}\right)^{12} = 10616.78}_{\text{compounded monthly at } r=0.06}.$$

The effective return rates for our previous example are included in Table 2.

**Example 2.6**

Suppose an investment $P$ grows at an APR of 5% compounded quarterly. Determine the number of years necessary for the investment to double in size.

*Solution.* We don't have a value for $P$, but this will turn out to be unnecessary. Otherwise, our variables are $r = 0.05$ and $t = 4$. We want the value of our investment to double, so we set $S = 2P$. Substituting this information into (2.5) gives

$$2P = P(1+0.05/4)^{4n} \quad \Rightarrow \quad 2 = (1.0125)^{4n}.$$

We're after the number of years $n$, which we can solve by taking logarithms:

$$\ln(2) = \ln(1.0125^{4n}) \quad \Leftrightarrow \quad \ln(2) = 4n\ln(1.0125) \quad \Leftrightarrow \quad \frac{\ln(2)}{4\ln(1.0125)} = n.$$

Evaluating $n$ gives the value $n = 13.94$ or approximately 14 years. ■

**Example 2.7**

An initial investment of \$5000 is made in January 2010, growing at an APR of 4% compounded semi-annually. In January 2015, a second investment of \$5000 is made into the same account. Determine the value of the account in January 2020.

*Solution.* Pretend for the moment that the second investment never occurs. The value of the investment made in January 2010 after 10 years is

$$S_1 = 5000\left(1+\frac{0.05}{2}\right)^{2\times 10} = 5000(1+0.025)^{20} = \$8193.08.$$

The value of the second investment spans a 5 year period, and hence is

$$S_2 = 5000\left(1+\frac{0.05}{2}\right)^{2\times 5} = 5000(1+0.025)^{10} = \$6400.42.$$

The total value of the investment in January 2020 is the sum of these (why?), yielding $S = S_1 + S_2 = 8193.08 + 6400.42 = \$14,593.50$. ■

### 2.1.3 Present Value

Our concern in Section 2.1 was determining the value of an account at some point in the future. There are cases when we need to move in the opposite direction; namely, we know the future value of an investment – or have a desired future value – and need to determine the corresponding present value. Luckily, this simply amounts to solving (2.5) for $P$ instead of $S$.

**Present Value I:** Suppose an investment is known to have a value of $S$ after a single compound period at a rate $r$. The *present value* of the investment is

$$P = \frac{S}{(1+r)} = S\,(1+r)^{-1}. \tag{2.7}$$

Again, this comes from the fact that after one compounding period, we know $S = P(1+r)$, wherein we then solve this equation for $P$. One paradigm for present values is as the amount of money you would need to invest now to guarantee a value $S$ in the future. Additionally, note that $P$ becomes smaller as $n$ becomes larger. Think about this for a moment: The further away the realization of a cash event, the less its worth in terms of current dollars.

We can then perform the same series of exercises as above.

**Example 2.8**

Suppose an investment is known to have a value of $S$ after $N$ compounding periods at a rate $r$. Determine the present value of this investment.

*Solution.* We can perform a similar exercise to determine the general expression for the present value. We'll see that the procedure amounts to multiplying by $(1+r)^{-1}$ for each compounding period that we move money back through time, but the conceptual hoops we must jump through are a little tricker than the future value.

If the investment $S$ occurs one compounding period into the future, we know that $P_1 = S(1+r)^{-1}$. If the investment is realized after two compounding periods, then the amount of money we need to invest today to achieve $S$ after two periods is the same as the amount of money we need to invest to achieve $P_1$ after one period. This is because we know that $P_1$ will evolve into $S$ after another compounding period. Thus

$$P_2 = P_1(1+r)^{-1} = S(1+r)^{-1}(1+r)^{-1} = S(1+r)^{-2}.$$

Continuing in this way, we will get $P_3 = S(1+r)^{-3}$, and so $P_N = S(1+r)^{-N}$.

A more direct solution is to recall that after $N$ compounding periods, the future value of our account is $S = P(1+r)^{-N}$. We're looking for the value of $P$, for which we solve to get

$$P = \frac{S}{(1+r)^N} = S(1+r)^{-N}. \qquad \blacksquare$$

Again, you'll likely need to invest some though into Example 2.8 to make sense of the reason why $P_2 = P_1(1+r)^{-1}$. However, the process of "moving money backwards through time" is to multiply by $(1+r)^{-1}$.

We will leave the following to Exercise 2-10 (which you should definitely work through), but if $S$ is the value of an investment $n$ years in the future, which we know grows at an annual percentage rate of $r$ compounded $t$-times yearly, the *present value* of the account is

$$P = S\left(1+\frac{r}{t}\right)^{-nt}. \tag{2.8}$$

**Example 2.9**

You wish to a purchase a \$20,000 car in 4 years as a graduation gift. To this end, you decide to invest some money in a Guaranteed Investment Certificate generating 3.2% per year compounded annually. How much money must you invest now to ensure you can purchase the car in 4 years.

*Solution.* Our variables are $S = 20000$, $t = 1$, $n = 4$, and $r = 0.032$. Substituting into (2.8) yields

$$P = 20000\,(1+0.032)^{-4} = \$17632.39. \qquad \blacksquare$$

Of course, this is not a great example since you're a poor student and that's a lot of money to sacrifice. We'll see a better approach in Example 2.16.

**Example 2.10**

Your cousin approaches you, boasting of an excellent investment opportunity. He is looking for a \$30,000 investment loan, guaranteeing cash payments in years 2, 3, and 4 of \$15,000, \$10,000 and \$8,000 respectively. You know that this money invested in a diversified equity portfolio will return 8% per year in the long run. Is it more profitable to invest in your cousin's business, or the market?

*Solution.* We compute the present value of the cash flows assuming they were to grow at 8% per year. Doing this we find

$$P = 15000(1.08)^{-2} + 10000(1.08)^{-3} + 8000(1.08)^{-4} = \$26,678.$$

This is less than our initial investment of \$30,000, showing that the present value of your returns in your cousin's company is less than your investment. You would therefore be better off investing your money in the equities portfolio. ■

The same visualization scheme that we used for future value (Figures 2.1 and 2.2) can be used with the present value. The difference is that now we move money to the left (backwards through time), and each pass through a compounding period incurs at $(1+r)^{-1}$ term.

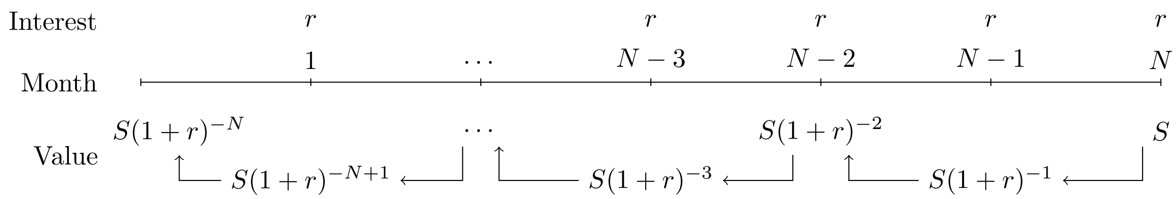



Figure 2.3: Moving money "backwards through time" incurs a $(1+r)^{-1}$ for each compounding period.

**Example 2.11**

Recall the compounding interest scheme of Example 2.2, wherein interest is compounded at a rate $r$ at the end of odd days of June, and at a rate $s$ at the end of even days of June. It's currently June 20th, and you need to cover your rent of $R$ dollars due at the beginning of the next month (after the June 30th). If you invest according to this scheme, how much money must you invest to cover your rent?

*Solution.* One way to attack this problem is to pull the money back through time using a diagram similar to Figure 2.3. Give this a try and ensure that you get the same answer below.

Another simple way of doing this is to modify Example 2.2. If we invested $P$ dollars at the beginning of day twenty, then by the end of day thirty the future value is

$$S = P(1+r)^5(1+r)^5.$$

The present value is then $P$, which we can solve for to get $P = S(1+r)^{-5}(1+s)^{-5}$. ■

### 2.1.4 Bonds

A bond is a debt, often belonging to a government or a corporation. You purchase a bond by giving the bond issuer cash, and in return the issuer agrees to pay back the original amount, plus periodic small payments designed to entice you into giving them the cash in the first place. The following words are used when describing bonds:

- **Maturity** - The period over which the debt is lent. At maturity, the debt is returned in full.
- **Face value** - The amount of the original debt, which is paid back to the purchaser upon maturity of the bond.

- **Coupon** - Small periodic payments made to the purchaser in order to incentivize the debt lending.

**Example 2.12**

A colleague of yours is trying to dump a 3 year bond whose face value is \$1000, matures in 1 year, and pays a coupon of \$50 every six months. You will receive the month 6 and 12 coupons, plus the principal. Determine the present value of the bond assuming the interest rate is quoted at an APR of 5% compounding

a) Semi-annually (every 6 months).

b) Monthly (every month).

*Solution.* That the original maturity of the bond was 3-years is a red-herring, and is not important to the problem.

a) Let's assume everything compounds semi-annually. The face-value will be paid out one year from now, or equivalently in two compounding periods. Thus the present value of the face-value payment is

$$P = 1000\left(1+\frac{0.05}{2}\right)^{-2\times 1} = \$951.84.$$

In this scenario, coupons happen to be paid at precisely the same frequency as the compounding period. We will receive the 6 month coupon after one compounding period, and the 12 month coupon after two compounding periods, thus the present values of these coupons are

$$C_1 = 50\left(1+\frac{0.05}{2}\right)^{-1} = \$48.78 \quad \text{(6 month coupon)}$$
$$C_2 = 50\left(1+\frac{0.05}{2}\right)^{-2} = \$47.59 \quad \text{(12 month coupon)}$$

The total present value is the sum of the present values of all payments, and so is

$$P + C_1 + C_2 = 951.84 + 48.78 + 47.59 = \$1048.21.$$

b) In the monthly compounding scheme, much of the mathematics is the same: We still compute the present value of the face-value and the two coupons. However, both the term-interest rate and the number of compounding periods have changed. The face-value is paid at 12 months, which occurs in twelve compound periods, so

$$P = 1000\left(1+\frac{0.05}{12}\right)^{-12} = \$951.33.$$

Notice the new term interest rate of 0.05/12. Similarly, the first coupon is paid after 6 compounding periods, and the second after 12 compounding periods, so

$$C_1 = 50\left(1+\frac{0.05}{12}\right)^{-6} = \$48.77 \quad \text{(6 month coupon)}$$
$$C_2 = 50\left(1+\frac{0.05}{12}\right)^{-12} = \$47.57 \quad \text{(12 month coupon)}$$

While these numbers are different from part (a), we notice that the difference only amounts to a few cents. The total present value is again the sum, giving

$$P + C_1 + C_2 = 951.33 + 48.77 + 47.57 = \$1047.67. \qquad \blacksquare$$

This idea of summing multiple present values is something we'll look at in greater detail in Section 2.3, but the basic principal is that these sorts of computations are additive.

## 2.2 Continuous Compounding Interest

Recall that a principal $P$ compounding $t$ times yearly at an APR of $r$ has an $n$-year future value of

$$S = P\left(1 + \frac{r}{t}\right)^{nt}.$$

Compounding yearly corresponds to $t = 1$, while daily is $t = 365$. We could, in theory, compound every second ($t = 31,557,600$), or every millisecond, etc. It's difficult to think about conceptually, but we'd like a notion of continuous compounding interest; that is, a scenario in which we compound at every instant without discrete breaks in time. We'll discuss the benefits of doing this in a moment.

If we want to compound "infinitely often," we might explore what happens to the above equation as $t$ becomes infinitely large. We don't have the tools to show this mathematically (yet), nor can I really explain what the following equation even means, but

$$\lim_{t\to\infty} P\left(1 + \frac{r}{t}\right)^{nt} = Pe^{rn},$$

where $e \approx 2.7187$ is Euler's number. This means that over $n$ years,the future value of a continuously compounding interest scheme can be computed using the natural exponent.

**Continuous Compounding Interest:** If a principal $P$ is invested at an annual percentage rate $r$, compounding continuously, then the future value of $P$ after $n$ years is

$$S = Pe^{rn}.$$

By the same token, the present value is

$$P = Se^{-rn}.$$

I've included the present value computation here as a simple application of the future value formula, where we solve for $P$. Let's try a few examples.

**Example 2.13**

Suppose $\$10,000$ is invested at a rate of 6% compounded continuously. Determine the value of the investment after 3 years.

*Solution.* This is an extension of Example 2.4 but now using continuously compounding interest. Indeed, the value of the account is

$$S = 10000e^{0.06\times 3} = 10000e^{0.18} = \$11972.20. \qquad \blacksquare$$

**Example 2.14**

Redo Example 2.12 but now assume that the market uses continuously compounding interest.

*Solution.* The present value of the \$1000 bond in 1 year is

$$S = 1000e^{-0.05\times 1} = 1000e^{-0.05} = \$951.23$$

The present value of the coupons, still worth \$50 each, are

$$C_1 = 50e^{-0.05\times 1} = \$47.56, \quad C_2 = 50e^{-0.05\times 0.5} = \$48.77$$

For a total present value of

$$S + C_1 + C_2 = 951.23 + 48.77 + 47.56 = \$1047.56. \qquad \blacksquare$$

Using the exponential function to compute such things is curious, for several reasons. First, note that nothing compounds continuously, so using the exponential to model anything is unrealistic. On the other hand, with modern computers and calculators, computing $e^{rn}$ is theoretically nicer to deal with, and simpler than its discrete counterpart $(1 + r/t)^{nt}$. This theoretical simplicity manifests in two ways: The expression $e^{rn}$ consists of two variables while $(1 + r/t)^{nt}$ consists of three variables, and $e^x$ is inverted by $\ln(x)$, while $(1 + 1/x)^x$ is difficult to invert. Moreover, we've seen that while compounding more often does lead to higher returns, the difference starts to become miniscule as $t$ becomes very large. Indeed, the effective annual rate of a 10% APR between compounding every day versus every second is

$$\begin{array}{l}\text{EAR-Daily/}\\ \text{365-times}\end{array} = 10.515\% \quad \text{and} \quad \begin{array}{l}\text{EAR-Secondly/}\\ \text{31557600-times}\end{array} = 10.517\%.$$

So while the latter compounds significantly more often, the difference between them only amounts to 0.002 of a single percent.

Just as we did with the effective annual rate, we can ask about a corresponding *effective continuous rate*. If you are quoted an APR of $r_t$ compounding $t$-times per year, the effective continuous rate is the value $r_c$ such that after one year, both your APR investment at at a rate $r_t$, and your continuous investment at a rate $r_c$ are the same; that is,

$$Pe^{r_c} = P\left(1 + \frac{r_t}{t}\right)^t$$

for some initial principal P. The value of $P$ quickly disappears, and we can solve this for $r_c$ to get

$$r_c = t\ln\left(1 + \frac{r_t}{t}\right). \tag{2.9}$$

For example, if we compound quarterly at 8%, the effective continuous rate is

$$r_c = 4\ln\left(1 + \frac{0.08}{4}\right) \approx 0.0792105....$$

So compounding quarterly at 8% will yield the same returns as compounding continuously at 7.92%, and you are free to check this by verifying that

$$e^{0.0792} \approx 1.0824 \approx \left(1 + \frac{0.08}{4}\right)^4.$$

The effective continuous rate means that we can convert discrete problems into continuous ones, and preserve the exact annual returns of the investment.

## 2.3 Annuities

Annuities represent the more realistic situation where an investment is supplemented at regular intervals. For example, when saving for your retirement, you may decide to invest \$1000/month into an investment account. Your money is now growing through the investment, and the principal is supplemented monthly. The amount of time between supplements is known as the *term*, so in the above example the term is monthly. There are two types of annuities: Those in which the payment is made at the end of the term – called *ordinary annuities* – and those where the payment is made at the beginning of the term – known as *annuities due*. Unless otherwise stated, we'll work with ordinary annuities.

There is no major change to determining the future and present values of an annuity: We compute the value of each possible payment, and move it either forward or backwards in time. The total sum of each principal then represents the total value of the investment. The major change in this section is that we will endeavour to use summation formulae to simplify our expressions.

### 2.3.1 Future Value of an Annuity

Let's begin by determining the future value of an annuity. Let's suppose that a regular recurring payment of $R$ dollars is made each term, we make $n$ regular payments, and the payments compound at a rate $r$ per compounding period.

There are two important time intervals here: The term between payments, and the compounding period. The easiest case to consider happens when these two coincide. At the end of the first term we make our first payment, so our account holds $S_1 = R$ dollars. At the end of the second month, we compound $S_1$ at $(1+r)$ and add another $R$ dollars, to get

$$S_2 = \underbrace{R(1+r)}_{S_1(1+r)} + R.$$

At the end of the third term, we compound $S_2$ at $(1+r)$ and make another payment of $R$ dollars, giving

$$S_3 = \underbrace{[R(1+r)+R](1+r)}_{S_2(1+r)} + R = R(1+r)^2 + R(1+r) + R.$$

We keep doing this, so that after $n$ terms our account holds

$$\begin{aligned} S &= R + R(1+r) + R(1+r)^2 + R(1+r)^3 + \cdots + R(1+r)^{n-1} \\ &= R\left[1 + (1+r) + (1+r)^2 + (1+r)^3 + \cdots + (1+r)^{n-1}\right] \\ &= R\sum_{k=0}^{n-1}(1+r)^k. \end{aligned}$$

This is a geometric series with common ratio $(1+r)$. Recall the closed form expression for a geometric series,

$$\sum_{k=0}^{n-1} c^k = \frac{1-c^n}{1-c},$$

into which we can substitute $c = 1 + r$ to get

$$S = R\sum_{k=0}^{n-1}(1+r)^k = R\frac{1-(1+r)^n}{1-(1+r)} = R\frac{(1+r)^n - 1}{r}.$$

**Future Value of an Annuity:** Suppose a recurring payment of $R$ dollars is invested at a rate $r\%$. If the payment schedule corresponds with the compounding term, the future value of the annuity after $N$-terms is

$$S = R\frac{(1+r)^N - 1}{r}. \tag{2.10}$$

It was essential to our derivation that the payment term and compounding terms coincide. When the two intervals do not agree, no attempt to finagle the formula will be successful. Instead, you must re-derive the formula from scratch.

Similar to simple compounding interest, if quoted a annual percentage rate of $r$ compounding $t$-times per year, the annuity formula (2.10) for an $n$-year period becomes

$$S = R\frac{(1+r/t)^{nt} - 1}{r/t}. \tag{2.11}$$

**Example 2.15**

In planning for your retirement, you invest \$2000/month at an annual percentage rate of 6% compounded monthly. Determine the value of your retirement savings after 20 years.

*Solution.* We use the future value equation, setting $R = 2000$, $r = 0.06/12 = 0.005$, and $n = 12 \times 20 = 240$. This gives

$$S = 2000\frac{(1+0.005)^{240} - 1}{0.005} = \$924,081.79. \qquad \blacksquare$$

**Example 2.16**

Reconsider the problem of buying yourself a graduation gift four years from now, in the form of a \$20,000 car. Investing monthly in a fund which returns an APR of 3.6% per year compounded monthly, determine your monthly payment.

*Solution.* We still use (2.10), but we set $S = 20000$, $r = 0.036/12 = 0.003$, and $n = 12 \times 4 = 48$. Substituting gives

$$20000 = R\left[\frac{1.003^{48} - 1}{0.003}\right] \approx 51.55R.$$

Solving for $R$ gives $R = \$388.01/\text{month}$. ■

**Example 2.17**

Suppose a regular recurring deposit of $R$ dollars is made into a High Interest Saving Account semi-annually. Interest is compounded at the end of each month at a rate of $r$. Determine the value of the account after $N$ years.

*Solution.* We will mimic our derivation of Equation (2.10), but this time we keep in mind that there are six compounding periods for each deposit period. As we've done in the past, it can be helpful to refer to a diagram to see how the payments work:

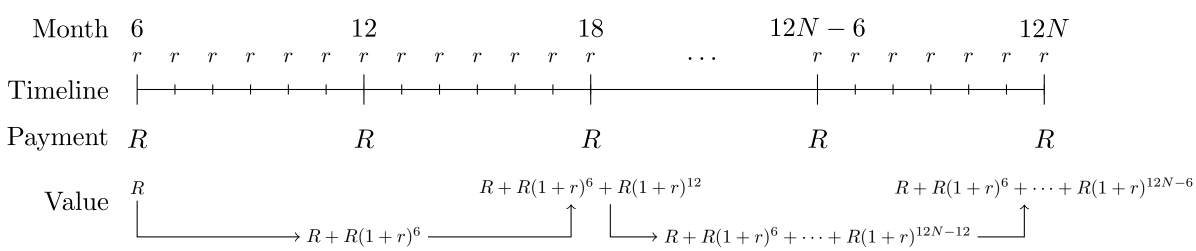



Figure 2.4: Compounding six-times more frequently than the payments. This diagram allows us to visualize how the annuity is constructed.

We make an initial deposit $S_6 = R$. Our next deposit occurs after 12 months, in which case the original deposit has accumulated six compounding periods, giving an account value of

$$S_{12} = R + R(1 + r)^6.$$

The next deposit occurs at the 18 month mark, wherein we accumulate $(1 + r)^6$ worth of interest on $S_{12}$, and add another $R$ to the account:

$$S_{18} = R + S_{12}(1 + r)^6 = R + \left[R + R(1 + r)^6\right](1 + r)^6 = R + R(1 + r)^6 + R(1 + r)^{12}.$$

Equivalently, the above formula represents the sum of each of the three payments pushed forward to the 18 month mark. If we continue this, we will finish with a balance of

$$\begin{aligned} S &= R + R(1+r)^6 + R(1+r)^{12} + \cdots R(1+r)^{12N-12} + R(1+r)^{12N-6} \\ &= R \sum_{k=0}^{2N-1} (1+r)^{6k}. \end{aligned}$$

Here we've identified that the common ratio between subsequent terms is $(1+r)^6$, and the $2N-1$ occurs because in $N$ years of semi-annual deposits, there are $2N$ payments.

Once again using our formula for the sum of a finite geometric sequence,

$$R \sum_{k=0}^{2N-1} (1+r)^{6k} = R \left[ \frac{1 - \left[(1+r)^6\right]^{2N}}{1-(1+r)^6} \right] = R \left[ \frac{1-(1+r)^{12N}}{1-(1+r)^6} \right].$$

There is no further simplification that is meaningful. ■

### 2.3.2 Present Value of an Annuity

Next we determine the present value of an annuity. In the case of simple compounding interest this was easy, as both the present and future values appeared in a single equation:

$$\underbrace{S}_{\text{Future}} = \underbrace{P}_{\text{Present}} (1+r)^N.$$

A similar phenomenon occurred with continuously compounding interest. The same cannot be said of the future value of an annuity,

$$A = R \left[ \frac{(1+r)^N - 1}{r} \right].$$

Here $A$ is the future value and $R$ is the *recurring payment amount*, nowhere does the present value make itself shown. Consequently, we'll have to compute the present value formula from scratch.

Like the future value, the present value is found by summing the present value of each payment. Again let's assume a recurring payment of $R$ dollars, with the ambient market rate at $r$. As before, both the compounding period and deposit term are important, with the easiest scenario arising when both coincide. We can use Figure 2.5 to help us reason through how the argument should go.

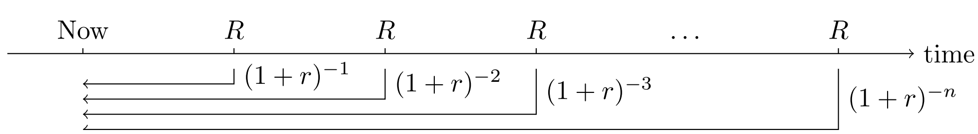



Figure 2.5: The present value of an annuity is determined by "bringing back" each future payment to the present.

The present value of the first payment is $R(1+r)^{-1}$, as indicated by the first arrow in Figure 2.5. Similarly, the present value of the second payment is $R(1+r)^{-2}$. This procedure continues until we arrive at the terminal payment, with a present value of $R(1+r)^{-n}$. Summing these together, we arrive at the total present value of the annuity

$$\begin{aligned} A &= R(1+r)^{-1} + R(1+r)^{-2} + R(1+r)^{-3} + \cdots + R(1+r)^{-n} \\ &= R\left[(1+r)^{-1} + (1+r)^{-2} + (1+r)^{-3} + \cdots + (1+r)^{-n}\right] \\ &= R\sum_{k=1}^{n}(1+r)^{-k}. \end{aligned}$$

As written, this is a geometric series whose common ratio is $(1+r)^{-1}$. Recalling the formula for the sum of a finite geometric series,

$$\sum_{k=1}^{n} c^k = \frac{c(1-c^n)}{1-c},$$

we substitute $c = (1+r)^{-1}$ and perform some algebraic manipulations to get

$$\begin{aligned} A = R\sum_{k=1}^{n}(1+r)^{-k} &= \frac{(1+r)^{-1}[1-(1+r)^{-n}]}{1-1/(1+r)} \\ &= R\left[(1+r)^{-1}[1-(1+r)^{-n}]\right]\frac{1+r}{r} \\ &= R\left[\frac{1-(1+r)^{-n}}{r}\right]. \end{aligned}$$

**Present Value of an Annuity:** Suppose a recurring payment of $R$ dollars is made, compounding at a rate $r\%$. If the payment schedule coincides with the compounding schedule, the present value of the annuity after $N$ terms is

$$A = R\left[\frac{1-(1+r)^{-N}}{r}\right]. \tag{2.12}$$

One can think of the present value of an annuity as the amount of money which needs to be invested now in order to cover all future payments. This viewpoint will be essential when we talk about amortization. If we're quoted a nominal rate of $r\%$ compounded $t$-times per year, Equation (2.12) becomes

$$A = R\left[\frac{1-(1+r/t)^{-nt}}{r/t}\right].$$

**Example 2.18**

You just won the Lotto 6/49, worth \$150 million dollars. You are presented with two options: The first is to receive an annual payment of 5 million dollars per year for the next 30 years. The alternative is to accept a one-time lump sum payment worth 90 million dollars. If you are confident you can get a yield return of 4% off any investment, determine which option is better.

*Solution.* We need to determine the present value of this annuity, which pays 5 million dollars per year. Using (2.12) with $R = 5000000$, $r = 0.04$, and $n = 30$, we get

$$A = 5000000 \left[ \frac{1 - (1 + 0.04)^{-30}}{0.04} \right] = \$86,460,166.50.$$

Thus the present value of the annuity is around \$86.5 million, which is less than the lump sum payment you're being offered. This suggests taking the lump sum payment. Alternatively, if you work for the lottery company, you must invest 86 million dollars now to cover the regular \$5 million payments to the winner. ■

**Example 2.19**

A bond with a face value of \$1000 and maturity of 5 years is sold, offering a 5% coupon quarterly. Determine the present value of the bond of the risk free interest rate is an APR of 5% compounded quarterly.

*Solution.* The present value of the bond itself, paying out \$1000 in 5 years, is

$$P = 1000 \left( 1 + \frac{0.05}{4} \right)^{-20} = \$780.01.$$

The coupons represent an annuity, paying \$50 quarterly, and have a present value of

$$A = 50 \left[ \frac{1 - (1 + 0.05/4)^{-20}}{0.05/4} \right] = \$879.97.$$

Thus the present value of the bond is $P + A = 780.01 + 879.97 = \$1659.98$. Compare this to the total payout of the bond, which is \$2000. ■

**Example 2.20**

Repeat Example 2.17 but now compute the present value of this annuity.

*Solution.* This is best represented by a hybrid of Figures 2.4 and 2.5, shown in Figure 2.6.

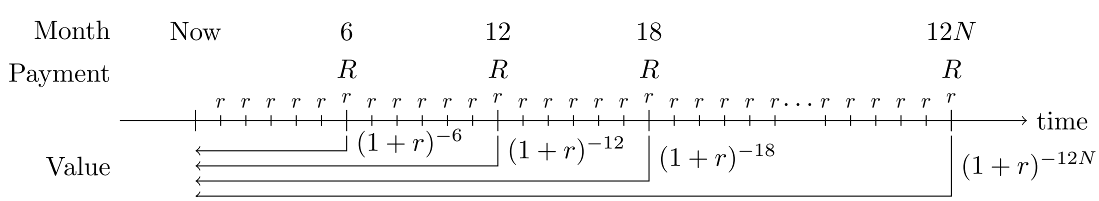



Figure 2.6: The present value of an annuity which compounds six times more frequently than the deposits.

Summing together the present values of each payment, we get

$$\begin{aligned} A &= R(1+r)^{-6} + R(1+r)^{-12} + R(1+r)^{-18} + \cdots + (1+r)^{-12N+6} + R(1+r)^{-12N} \\ &= \sum_{k=1}^{2N} (1+r)^{-6k}. \end{aligned}$$

Using our expression for the sum of a finite geometric series with common ratio $(1+r)^{-6}$ we get

$$A = R(1+r)^{-6} \frac{1 - \left[(1+r)^{-6}\right]^{2N}}{1-(1+r)^{-6}} = R\frac{1-(1+r)^{-12N}}{(1+r)^6 - 1}. \qquad \blacksquare$$

### 2.3.3 Amortization

Amortization is the process of paying off a loan in regular, uniform payments. For example, car payments and mortgage payments are made through amortization. Because such payments are regular and uniform, they are a form of annuity, albeit we have to think about them in a slightly different way.

The way to think about loans is that *the bank is purchasing an annuity from you.* To do this, the bank is going to give you a lump sum payment $A$ today, and in return you will pay to the bank regular payments of $R$ dollars. The value of this annuity is precisely the amount of money the bank gives to you, and this is the present value since it's given to you today. Hence amortization is calculated using the present value of an annuity formula.

Suppose you work for a bank and a client applies for a $\$10,000$ loan at an annual percentage rate of 4%. The client will make monthly payments and hopes to have this debt eliminated after 2 years. The present value of this annuity is precisely $\$10,000$, so we can substitute everything into (2.12) to find

$$10000 = R\left[\frac{1-(1+0.04/12)^{-24}}{0.04/12}\right] \approx 23.03R$$

which we can solve for $R$ to find $R = \$434.25$. Alternatively, perhaps the client decides they can pay the loan off at \$800/month. Substituting everything gives

$$10000 = 800\left[\frac{1-(1+0.04/12)^{-n}}{0.04/12}\right]$$

which we can solve for $n$ to get $n = 12.8$, so it will take approximately 13 months to pay off the loan with \$800 payments.

**Example 2.21**

You purchase a $\$800,000$ home with a 20% down payment. The bank approves you for a loan at an APR of 2.64% compounded monthly.[a]

1. Determine the monthly payments over a 25 year amortization.
2. If you can contribute \$5000/month to your mortgage, how many years will it take for you to pay off your home?

[a] In reality, you would never get this low a rate for a full 25 year mortgage. Rates are renegotiated regularly.

*Solution.* Make a 20% down payment means the value of the loan is $A = \$640,000$, with an interest rate of $r = 0.0264$. In both cases, we use the present value of an annuity formula (2.12).

1. We are solving for $R$ with the knowledge that $n = 25$, so substituting everything gives
$$640000 = R\left[\frac{1-(1+0.0264/12)^{-25\times 12}}{0.0264/12}\right] \approx 219.44R$$
which we solve for $R$ to find $R = \$2916.51$. Hence you must make a mortgage payment of \$2916.51 per month.

2. Now we have $R$ and wish to determine $n$. The present value of an annuity formula becomes
$$640000 = 5000\left[\frac{1-(1+0.0264/12)^{-n}}{0.0264/12}\right]$$
where $n$ is written in months. We could write the exponent as $-12n$, in which case $n$ would then be years. Solving this equation for $n$, we get $n = 150.5$ months, or just over 12 years. ∎

Summarizing the above information, the monthly payment of an amortization can be computed via the formula
$$R = A\left[\frac{r}{1-(1+r)^{-n}}\right] \tag{2.13}$$
while the number of payment installations required can be determined by solving this for $n$:
$$n = \frac{\ln(R)-\ln(R-Ar)}{\ln(1+r)}.$$
At the beginning of the $k$th payment period, the outstanding principal on the loan is the present value of the remaining payments. At the beginning of the $k$th period, there are $n-k+1$ remaining payments, so
$$\begin{array}{r}\text{Principal outstanding at the}\\ \text{beginning of } k\text{th payment}\end{array} = R\left[\frac{1-(1+r)^{-n+k-1}}{r}\right]$$
The interest paid on on the $k$th payment is $r$ times the principal outstanding at the beginning of the $k$th payment, so
$$\text{Interest in the } k\text{th payment } = Rr\left[\frac{1-(1+r)^{-n+k-1}}{r}\right] = R\left[1-(1+r)^{-n+k-1}\right].$$
The principal paid in the $k$th payment is the payment $R$ less the interest paid in the $k$th payment, so
$$\text{Principal covered in } k\text{th payment } = R - R\left[1-(1+r)^{-n+k-1}\right] = R(1+r)^{-n+k+1}.$$
The total interest is the sum of all the payment $Rn$ less the value of the loan $A$,
$$\text{Total Interest Paid } = nR - A.$$

**Example 2.22**

Determine, as a fraction of the loan amount, the total amount of interest paid on any 25 year mortgage as a function of the interest rate $r$. Explicitly compute this fraction when the interest rate is an APR of 2.5%, 3%, and 3.5% monthly.

*Solution.* The total interest is $nR - A$. Here $n = 25 \times 12 = 300$, and we know $R$ from (2.13). Putting this together we get

$$I = nR - A = 300A\left[\frac{r/12}{1-(1+r/12)^{-300}}\right] - A = \left[300\left(\frac{r/12}{1-(1+r/12)^{-300}}\right) - 1\right]A.$$

The value in square brackets is the total interest as a fraction of $A$. For the different values of $r$ given, we get

| $r$ | 0.025 | 0.030 | 0.035 |
|---|---|---|---|
| $I$ | $\approx 0.35$ | $\approx 0.42$ | $\approx 0.50$ |

Here we see the impact of a half-percentage point. Suppose your mortgage loan is for $\$500,000$. At a 2.5% rate, you will pay $\$175,000$ in interest; at 3.0% you will pay $\$210,000$ in interest, and at 3.5% you will pay $\$250,000$ in interest. You can mitigate this gap by paying off your mortgage faster. For example, over a 15 year amortization schedule, the values for $I$ are $(0.20, 0.24, 0.29)$. Here the spread between 2.5% and 3.5% is only 0.09, as compared to 0.15 over a 25-year amortization. Hence shortening your schedule decreases your exposure to high rates. ■

**Example 2.23**

You have taken out a $\$500,000$ mortgage financed at an APR of 3% compounded monthly over 25 years. After 10 years, a wealthy great aunt dies and leaves you an inheritance of $\$200,000$. You decide to refinance your mortgage by contributing the entire inheritance towards the principal. Assuming you keep the same mortgage terms (rate, compounding period, original maturity), determine your monthly payments on the newly financed mortgage.

*Solution.* Our first step is to determine the monthly payments on the home under the original mortgage. Here $A = 500000, r = 0.02, n = 25, t = 12$ so

$$R = 500000\left[\frac{0.03/12}{1-(1+0.03/12)^{-12\times 25}}\right] \approx \$2371.06$$

each month.

The next step is determine the principal remaining on the mortgage after 10 years, which is the present value of the annuity at that instance. The first payment after 10 years of compounding monthly is the 121st period, so the remaining balance on the home is

$$A' = 2371.06\left[\frac{1-(1+0.03/12)^{-180}}{0.03/12}\right] = 343,342.46.$$

You now apply the $\$200,000$ inheritance to this, bringing it down to $\$143,342.46$, and refinance at 3% amortized over the remaining 15 years:

$$R = 143342.46\left[\frac{0.03/12}{1-(1+0.03/12)^{-12\times 15}}\right] \approx \$989.90.$$

Thus your monthly payments become $\$989.90$. ■

## 2.4 Perpetuities

Perpetuities are annuities with no terminal date – they pay out forever. For example, the university might choose to start a scholarship bestowing \$1000/year every year without end. To fund this scholarship, the university deposits a lump sum payment into an investment account, with the understanding that \$1000 will be withdrawn each year to cover the scholarship. The goal then is to ensure that the return on investment is enough to cover the scholarship in perpetuity; that is, forever. How much money should the university deposit if they conclude a long-term return of $r$ per year?

This sounds a lot like the present value of an annuity, with the caveat that the series extends forever. Imitating that derivation, let's suppose that regular payments $R$ are to be made at a rate of $r$, and that the compounding period coincides with the deposit period. The present value of the perpetuity is identical to that presented at the beginning of Section 2.3.2, but now forms an infinite series

$$A = R(1+r)^{-1} + R(1+r)^{-2} + R(1+r)^{-3} + \cdots = \sum_{k=1}^{\infty} R(1+r)^{-k}.$$

The formula for the sum of an infinite geometric series is

$$\sum_{k=0}^{\infty} c^k = \frac{1}{1-c} \qquad \text{so} \qquad \sum_{k=1}^{\infty} c^k = \frac{1}{1-c} - 1 = \frac{c}{1-c},$$

into which we substitute $c = (1+r)^{-1}$ to get

$$A = R\sum_{k=1}^{\infty}(1+r)^{-k} = R\frac{(1+r)^{-1}}{1-(1+r)^{-1}} = R\frac{(1+r)^{-1}}{r(1+r)^{-1}} = \frac{R}{r}.$$

**Present Value of a Perpetuity:** Suppose a recurring payment of $R$ dollars, compounding at a rate $r$. If the payment schedule coincides with the compounding schedule, the present value of a perpetuity is

$$A = \frac{R}{r} \tag{2.14}$$

**Example 2.24**

Suppose the university is establishing a scholarship to pay \$1000/year. To do this, the institution must deposit a lump sum money into an investment account, which it figures will grow at a rate of 4% per year. Determine how much must be invested for the scholarship to run forever.

*Solution.* We know $R = 1000$ is to be paid annually, and $r = 0.04$. The present value of the perpetuity is the amount the university must invest, which we find to be

$$A = \frac{R}{r} = \frac{1000}{0.04} = \$25,000.$$ ■

**Example 2.25**

Repeat Example 2.20 wherein the compounding period occurs six times more frequently than the deposit period, but now assume that the deposits happen in perpetuity. Determine the present value of that perpetuity.

*Solution.* We can steal the answer directly from the solution to Example 2.20, and covert it into a perpetuity. In this case we get

$$A = R(1+r)^{-6} + R(1+r)^{-12} + R(1+r)^{-18} + \cdots = \sum_{k=1}^{\infty}(1+r)^{-6k}.$$

Again using the formula for an infinite geometric series, we can reevaluate this expression as

$$A = \frac{(1+r)^{-6}}{1-(1+r)^{-6}} = \frac{1}{(1+r)^6 - 1}. \qquad \blacksquare$$

## 2.5 Exercises

2-1. The Wacky Fund™ compounds interest at the end of each weekday (MTWRF). Each situation below describes how interest is computed. Determine the future value of the fund after the prescribed amount of time, given that you invest your money on Sunday.

| Situation | M | T | W | R | F | Time | Notes |
|---|---|---|---|---|---|---|---|
| I | $r$ | $r$ | $r$ | $r$ | $r$ | 2 weeks | $r > 0$ |
| II | $r$ | $s$ | $r$ | $s$ | $t$ | 3 weeks | $r, s, t > 0$ |
| III | $r$ | $2r$ | $3r$ | $4r$ | $5r$ | 10 days | $r > 0$ |
| IV | $2r$ | $r$ | $s$ | $2s$ | $t$ | 15 days | $r, s, t > 0$ |

2-2. On January 1st you invest $P$ dollars into two different bank accounts. The first account compounds interest at a term rate of $r$ at the end of odd months (January, March, ...). The second account compounds interest at a rate of $s$ at the end of even months (February, April, ...). What are the values of these accounts after 1 year?

2-3. Suppose you invest in a bank account which compounds at an interest rate $r$ at the end of each month. On January 1st you invest $P$ dollars. On February 1st (after the first compounding period) you withdraw $P/2$ dollars, on March 1st you withdraw $P/4$ dollars, and on April 1st you withdraw $P/8$ dollars. What's the value of the account on May 1st?

2-4. Suppose you invest in a bank account which compounds interest at a rate $r$ at the end of each month. On January 1st, you invest a value of $P$ dollars into the account, and continue to invest $P$ dollars on the first day of each month for a year. What's the value of the account on January 1st of the next year?

2-5. In each situation described below, identify the initial payment, the term interest rate, and the number of compounding periods. Use this to express the future value of the asset. Some situations may involve multiple assets, in which case you should extract the above information for each asset.

(a) Your uncle Vernon invests \$200 compounding with an APR of 1% every 6 months, and he does this for five years.

(b) Your aunt Petunia invests \$1000 in a high interest savings account at an APR of 2.3%, compounded daily for 3 years.

(c) Your cousin Dudley puts \$500 into a tax free savings account on January 1st, 2015 and January 1st, 2017. The investment compounds monthly at 1.8%, and it is now September 1, 2019.

(d) Your uncle's sister Marjorie (to whom you are not related) is a loan shark. Your friend Ron borrowed \$1200 from her 63 days ago. She charges him 5% every 9 days, and he hasn't made any payments.

2-6. Determine the value of each investment

(a) An investment of \$10000 at an APR of 5.2% compounded annually for 13 years.

(b) An investment of \$5000 at an APR of 3% compounded monthly, followed by another investment of \$5000 made 2 years after the first. What is the value of the account after 5 years?

(c) Two investments, one of \$2000 at an APR of 2% compounded yearly, and one of \$1000 at an APR of 2.5% compounded monthly, after 20 years.

(d) An investment of \$10000 made in January 2018 at an APR of 5% compounded monthly, and another investment of \$20000 made in January 2025 at an APR of 3% compounded weekly. What is the value of the account in January 2030?

2-7. In each situation, choose the situation which is better for you.

(a) You are late on a bill payment of \$10000. The contractor offers you two options: To have a late fee applied at an APR of 10% compounded monthly, or an APR of 9.8% compounded daily. You will pay the bill in exactly one month.

(b) You are hurt in an accident and the insurance company offers you two payout schemes. You can take a payout now of \$30000, or three payouts of \$12000, one now, one after 2 years, and one after 4 years. Suppose the risk free interest rate is 7%.

2-8. In each case below, you want to invest money and are presented with two investment schemes. Determine which option is better (has a higher effective rate).

(a) Option A: 3% APR compounding monthly. Option B: 2.8% daily.

(b) Option A: 2.5% APR compounding weekly. Option B: 2.4% monthly.

(c) Option A: 5.2% APR compounding yearly. Option B: 5% weekly.

(d) Option A: 10% APR compounding yearly. Option B: 12% biweekly.

(e) Option A: 1% APR compounding daily. Option B: 0.9% continuously.

2-9. In each case, determine which invest yields the greater returns.

(a) Option A: \$1000 invested at an APR of 8% compounding biweekly over 2 years.

Option B: \$500 invested at an APR of 6% compounding monthly over 15 years.

(b) Option A: \$100 invested at an APR of 4% compounding continuously over 3 years.
Option B: \$100 invested at an APR of 5% compounding semi-annually over 1 year.

(c) Option A: \$50 in 6 months and \$50 in 12 months, assuming a risk-free rate of 3% compounding monthly.
Option B: \$95 today.

2-10. Derive Equation (2.8) by imitating the argument from Equation (2.5).

2-11. For each annual interest rate prescribed, determine the amount of time it takes for a principal $P$ to

i. double

ii. triple

iii. increase 10-fold

if compounded monthly at

(a) 4%

(b) 6%

(c) 8%.

2-12. Suppose an investment compounds at 5% APR monthly. Find the *effective daily rate*; namely, the annual percentage rate $r_d$ such that compounding at $r_d$ daily returns the same value after one year.

2-13. In question 1, is $r_a > r_d, r_a < r_d$, or is $r_a = r_d$. Explain without computing $r_a$.

2-14. Generalize Problem (1). Suppose an investment compounds at an APR of $r$, $t$-times per year. What is the effective $s$-rate; that is, the annual percentage rate $r_s$ such that compounding at an APR of $r_s$, $s$-times per year returns the same value?

2-15. Suppose you compound at a rate of 1% every 5 days. What the effective annual rate of this investment?

2-16. Suppose you compound using two schemes: At a rate $r_1 = 1/52$ every week, and at a rate $r_2 = 1/365$ every day (these are not APRs). Which of these schemes earns more money?

2-17. Your "T-Rex Score" is the percentage of an investment return you get to keep after fund fees have been subtracted; that is

$$\text{T-Rex Score} = \frac{\text{Net Return on Investment After Fees}}{\text{Total Return on Investment}}$$

This is typically calculated over a 25 year period. Common mutual funds have Management Expense Ratios (MERs, also known as fund fees) of 2%. Suppose that you realize an effective annual return of 8% every year. What is your T-Rex Score?

2-18. Consider a bond purchased January 2018 with face value \$1000 expiring in January 2020. This bond pays a coupon of \$30 every year on December 31. If the interest rate on the bond is 2% compounded annually, determine the current value of the bond.

2-19. Determine the amount of money you must invest today at an APR of 5% in order to have \$400,000 in 20 years. What is the corresponding number if the APR is changed to 4%? What about 6%? Assume you're compounding monthly.

2-20. Suppose `UTM` issues bonds with a one-year maturity and \$1000 face value on January 1, 2019. These bonds pay quarterly coupons of \$50 on April 1, 2019; July 1, 2019; October 1, 2019; and January 1, 2020.

(a) Suppose we're computing the present value of this bond using two different scenarios. In the first scenario, we assume an APR of 1.3%, and in the second scenario we assume an APR of 2%. Without doing any computations, which scenario results in a *lower* present value?

(b) Confirm your guess in part (a) by computing the present value in each scenario, assuming the given APR compounds monthly.

(c) Assume instead that the compounding periods occur quarterly, and coincide with the coupon payments. Modify your calculation in part (a) to compute the present value of the 1.3% APR bond.

(d) Again assume quarterly compounding periods and an APR of 1.3%. This time use the present value of an annuity formula to confirm the answer you got in part (c).

2-21. Consider the effective annual rate for an APR of $r\%$, described as $E(n) = \left(1 + \frac{r}{n}\right)^n - 1$.

(a) Does this number increase, decrease, or stay constant if $n$ increases?

(b) If you are investing, would your rather have discretely compounding interest or continuously compounding interest?

(c) If your interest rate is instead quoted as compounding every 2-years, what would $n$ be? What about every 10 years?

2-22. Suppose a principal $P$ is invested at an APR of $r\%$ compounded continuously over $n$ years. Indicate whether each statement is true or false, and justify your answer.

(a) It takes $\ln(3)/r$ years for $P$ to triple in size.

(b) If $P$ triples in 5 years, then $r = \sqrt[n]{e^5}$.

(c) The amount of time it takes for $P$ to triple depends on $P$.

(d) It will take longer for this investment to triple than if it were compounded monthly.

(e) None of the above statements are correct.

2-23. For each $r$ and $t$ below, determine the effective yearly rate, and the effective continuous yearly rate, on a principal growing at $r\%$ and compounding $t$-times per year.

(a) $r = 0.06, t = 2$

(b) $r = 0.04, t = 12$

(c) $r = 0.08, t = 52$

(d) $r = 1.00, t = 525,600$

Let $E$ be the effective annual rate, and $E_c$ be the effective continuous yearly rate. Is $E < E_c$ or $E > E_c$. Explain why.

2-24. Aziz and Bo are planning a big graduation party for their friends in exactly four years. They decide to invest some money today to cover the costs of the party. Both determine they will receive an APR of 3% from their investments and need \$10 000 to cover the costs, but Aziz's investment compounds monthly while Bo's compounds weekly.

If $P_A$ is the initial value of Aziz's investment, $P_B$ is the initial value of Bo's investment, and both aim to finance the party with the least amount of capital, what is the relationship between $P_A$ and $P_B$? Namely, is $P_A > P_B$, $P_B > P_A$, are they equal, or can they not be compared? Give a brief explanation of your answer.

2-25. Suppose the bank quotes you a loan with an APR of 3.2% compounded monthly. Determine the corresponding APR if the loan were instead compounded semi-annually, accurate to four decimal places.

2-26. It's January 1st. You are given the option of receiving \$1000 today, or \$510 at the end of June and 520 at the end of December. If you assume a market rate of 4% APR compounding monthly, which option is better?

2-27. An investment scheme compounds at a rate $r$ at the end of months that end in "uary", $s$ in months that end in "ber", and $t$ in all other months. You will invest $P$ dollars on January 1st 2020, and hope to have $A$ dollars on January 1st 2021. How much should you invest to day to make this happen?

2-28. Supppose you lend a friend \$100 on January 1st, and in return she offers to pay you back \$10 at the start of each month (including this month, for some weird reason) for one year (the last payment is January 1st of the next year). What is the present value of this deal if you assume a market rate of 5% APR compounding monthly? Is the present value worth more than the \$100 you leant her? (You don't have to use a calculator for this problem, but you might want to).

2-29. On January 1st, 2019 you put \$1000 into an account growing at 2% APR compounded daily. Starting that same day, you invest \$10/week into an account paying 3% compounded weekly. Write down the formula for the total value of your investments exactly one year (365 days) later.

2-30. Determine the value of each described annuity. Each annuity is ordinary, unless otherwise specified.

(a) Invest \$1000/month for 5 years, at a nominal rate of 3.2% compounding monthly.

(b) An initial lump sum payment of \$50,000, followed by semi-annual contributions of \$10,000, growing at a nominal rate of 5.7% compounding semi-annually, for 10 years.

(c) Invest \$1,000 monthly into a bond fund, with a 2.4% APR, compounded monthly. Invest another \$1,000 monthly into an equity fund returning 6.1% yearly, compounded monthly. Make both investments over a span of 20 years.

(d) Consider the same investment scheme as (c), but add in one time initial lump sum investment of \$100,000, split 20-80 into bond and equities.

2-31. On January 1st, 2017, your great uncle creates an account in your name and invests \$20,000 in treasury bills. The nominal rate on the T-bills is 1.95% compounded every 73 days. Exactly 6 months later, your mother deposits \$10,000 into an investment account, compounding continuously at 4% per year. What is the combined value of these two accounts on January 1st, 2019?

2-32. Suppose you have purchased a bond with face value \$1000. The bond pays \$20 coupons quarterly and expires in one year. Assume the risk free rate is an APR of 2%.

(a) Suppose we assume quarterly compounding periods. Express the present value of the bond in terms of an *annuity*.

(b) If we assume a weekly compounding period (13 weeks per quarter), find the present value of the bond.

2-33. It's January 1st, 2019 and you purchase a bond whose face value is \$1000 and pays out coupons of \$50. This bond pays its first coupon three months from now, and continues to pay coupons every quarter after that. The bond matures on January 1st, 2020.

(a) Write down a formula for the present value of the bond if the nominal rate of return is 3%, compounded monthly.

(b) Write down a formula for the present value of the bond if the nominal rate of return is 3%, compounded continuously.

2-34. Suppose $r, s, t > 0$ are interest rates, and compounding interest is applied at the end of each month according to the following scheme: At the end of the first month we compound with an interest rate $r$, at the end of the second month we compound with an interest rate $s$, and at the end of the third month we compound with an interest rate $t$. After this, the cycle begins again.

(a) If $R$ dollars is invested at the beginning of the first year, determine the return on investment after two years.

(b) You are saving to make a payment of $S$ dollars the end of August in the second year. Determine the amount of money that you must put away on January 1st of the first year, using the above compounding scheme, to ensure that you can cover this payment.

(c) Today is January 1st of the first year, and you have taken out a loan of $L$ dollars, amortized over the next three months. Interest is applied according to the above interest scheme. At the end of January, February, and March, you must make payments on this loan. What are your payments? Write your answer in terms of $L, r, s$, and $t$.

(d) Now suppose you amortized the same loan of $L$ dollars over three years, and are making payments of $A$ dollars monthly. What is the principal remaining on the loan at the beginning of August in the third year? Write your answer in terms of $L, A, r, s$, and $t$.

2-35. (a) On January 1st, 2018 `UTM` issues a 10 year corporate bond with a face value of \$1000. The prevailing interest rate means that \$10 coupons are paid out monthly. Determine the present value of the bond if the risk-free rate is 2.8% (assume monthly compounding interest).

(b) Five years later on January 1st, 2023 `UTSC` issues a 5 year corporate bond with a face value of \$800. However, interest rates have increased, so that this bond is paying monthly \$12 coupons. On this date, which of the `UTM` and `UTSC` bonds has a greater value? (Assume the risk free interest rate has not changed.)

(c) Generally, when interest rates (the value of the coupon) rises, what should happen to the price of any existing bond? Justify your response.

2-36. You invest $P$ dollars at an annual percentage rate of $r\%$ compounded continuously. How long does it take for your account to have $10P$ dollars?

2-37. Suppose a deposit of $R$ dollars is made at the end of each year, and assume an APR of $r\%$ compounding *continuously*.

(a) If we invest like this for 3 years, what is the value of the account after the third year?

(b) If we are making these payments for 3 years, how much money must we invest today to cover the payments?

(c) If we make these payments in perpetuity, find a formula which describes the amount of principal we must invest today to cover all future payments.

(d) If we invest like this for three years, but instead use an account with an APR $s\%$ compounded weekly, determine $s$ so that this second account has the same value as your answer in part (a) after three years.

2-38. Compute the *effective* continuous rate for an investment scheme compounding at 5% APR monthly.

2-39. If one uses continuous compounding interest to model the scheme above, determine the error in the model after 100 months. (Desmos Model)

2-40. Suppose a principal $P$ compounds continuously at a rate $r$ for one year, then compounds continously at a rate $s$ for one year. What is the value of the account after those two years?

2-41. Suppose an investment compounds continuous at a rate $r$. You make investments at the beginning of every week for 8 weeks. What is the value of the account after those 8 weeks?

2-42. Suppose an ordinary annuity consists of \$100 deposits made every week at an APR of 3%, compounding weekly. Use sigma notation to write down the future value of the annuity after 1 year.

2-43. You work for the shipping company Canada $e^{\text{x(press)}}$, and host a fleet of delivery vehicles. In 5 years the fleet will need an overhaul, which your team estimates at \$1, 500, 000. You set up a fund which pays 3.8% compounded monthly. Determine the monthly payments into this fund to ensure you can overhaul the fleet in 5 years.

2-44. In saving for your retirement, you figure you need about \$1, 000, 000 to maintain your lifestyle. You are currently 22, and figure you can safely put away \$1000/month into an account which averages 6.6% per year (compounded monthly). At what age can you retire?

2-45. You've won a \$20, 000, 000 lottery, and have the option of choosing an annual payment of \$1, 000, 000 for 20 years, or a lump sum payment of \$10, 000, 000 today. You figure you can make 7.7% in the equity market. Which option should you choose?

2-46. Suppose $r > 0$ describes an interest rate, and compounding interest is made according to the following scheme: After the first month we compound at an interest rate $r$, and after the third month we compound at an interest rate of $r$. After this, the cycle begins again.

(a) If $R$ dollars is invested at the beginning of the year according to the above compounding scheme, determine the return after four years.

(b) Determine the Annual Percentage Rate $s > 0$ such that if we compounded monthly at an APR of $s$, we would have the same return as the investment scheme above after one year.

(c) If interest is applied according to the scheme above, how much money must you invest today to ensure that your investment is worth $G$ dollars after four months?

(d) If a perpetuity of $R$ dollars is paid out every six months using the above compounding interest scheme, the present value of the perpetuity is

$$A = R(1+r)^{-4} + R(1+r)^{-8} + R(1+r)^{-12} + \cdots .$$

Use the infinite geometric series formula to evaluate this series, and simplify your answer as much as possible.

2-47. One retirement vehicle is to purchase an annuity through your bank, in which you deposit a certain amount of money up front, and the bank makes regular payments to you based off that deposit. Suppose you purchase such an annuity for $\$300,000$ at a nominal rate of $6.3\%$. Determine the expected monthly payments, if payments are guaranteed for 10 years.

*Note:* Often this is offered as part of a life insurance policy, where your retirement income is guaranteed for life. You have the additional option of guaranteeing a number of years, so that if you die before that period expires, the estate continues to claim the income. This is much harder to simulate, since it requires actuarial data.

2-48. Let $r > 0$ be an interest rate, and consider the following three compounding investment schemes:

**I:** Interest is compounded at a rate of $r$ at the end of every month.

**II:** Interest is compounded at a rate of $3r$ at the end of every third month.

**III:** Interest is compounded at a rate of $6r$ at the end of every sixth month.

(a) Suppose you invest $R$ dollars on January 1st. Rank these three schemes in terms of which provides you the greatest return after one year. For example, writing (II $>$ I $>$ III) means Scheme II returned the greatest return, while Scheme III provided the least return. In three sentences or less, explain why these are not equal, and why the top ranked scheme is the largest.

(b) Suppose you are investing according to Scheme III's compounding rate. What is the effective rate $r_2$ of Scheme II relative to Scheme III? That is, determine the rate $r_2$ such that using Scheme II provides the same return as Scheme III after one year.

(c) Let $r_{\mathrm{I}}, r_{\mathrm{II}}, r_{\mathrm{III}}$ denote the effective annual rates of each of the schemes above. Put these three numbers in ascending order. For example, $r_{\mathrm{II}} < r_{\mathrm{I}} < r_{\mathrm{III}}$. In three sentences or less, explain the ordering.

(d) Suppose perpetuities are to be paid according to each scheme above, with the payments coinciding with the compounding periods. Determine the present values $A_{\mathrm{I}}, A_{\mathrm{II}}, A_{\mathrm{III}}$ of each scheme.

(e) If we pay out a yearly perpetuity under Scheme III, the present value of this perpetuity is

$$A = R(1+6r)^{-2} + R(1+6r)^{-4} + R(1+6r)^{-6} + \cdots .$$

Use the infinite geometric series formula to evaluate this series, and simplify your answer as much as possible.

2-49. In planning for your retirement, you've concluded that you need \$30,000/year to live off of. You've established an investment portfolio that you're certain will return 5% long term. Determine the amount of money you need to retire.

2-50. Two investments are made at the same time. Investment A has a one-time deposit of \$1000 invested at an annual percentage rate of 5%, compounded yearly for 10 years. Investment B deposits \$100 into the account at the beginning of each year, and it grows at 5% compounded yearly. Let $V_A$ be the value of Investment A after 10 years, and $V_B$ be the value of Investment B after 10 years. Which statement is correct?

2-51. Consider a single lump sum of \$100,000 invested today at an annual percentage rate of 3% compounded quarterly. An additional sum of \$2000 is invested monthly thereafter at 5% compounded monthly. Determine the total value of the account after 10 years. Round to the nearest cent.

2-52. Consider a home you purchase for \$600,000, amortized over 25 years at a rate of 4.5%. Determine your monthly payments, and the total interest paid on the loan, if your down payment is

(a) 10% (b) 15% (c) 20%

2-53. If a loan amount of $R$ dollars is amortized with monthly payments at a nominal interest rate $r$%, write down an expression which indicates the number of months required to pay back the loan.

2-54. Consider a home you purchase for \$600,000, at a down payment of 15% and at a rate of 4.5%. Determine your monthly payments, and the total interest paid on the loan, if your amortization period is

(a) 15 years (b) 20 years (c) 25 years

2-55. Your rich aunt Hannelore is setting up a fund for victims of math tests. The fund will pay \$10 000 per year without end. If her financial advisors think they can get a 2% return (compounded yearly), how much money must be invested today to ensure the fund lasts forever?

2-56. The winner of the Cash-4-Ever Lottery receives \$1000 per week in perpetuity. For this question, assume an Annual Percentage Rate (APR) of 2%.

(a) In three sentences or less, explain how this company can pay \$1000 per week forever using only a finite initial principal.
(b) Derive the formula for the present value of the $k$th payment, where $k \geq 1$ is a natural number. *Note:* Your answer should depend on $k$.
(c) Write a formula for the present value of the Cash-4-Ever prize.
(d) Let PV($r$) be the present value of the Cash-4-Ever Lottery prize at an APR of $r$%. Suppose $r_1$ and $r_2$ are positive numbers with $r_1 > r_2$. Is PV($r_1$) greater than PV($r_2$)? Explain your reasoning in three sentences or less.

2-57. Suppose you aim to establish a perpetuity wherein money will be withdrawn every other compounding period. For example, the investment might compound every 6 months, but funds are withdrawn once a year.

(a) If $R$ dollars are withdrawn at a rate of $r\%$, find a formula for the present value of the perpetuity. Be sure to simplify your formula as much as possible.

(b) Suppose $P_1$ is the present value of an annuity whose compounding term corresponds with the withdrawal term. Let $P_2$ denote the present value of the situation described above, wherein the investment compounds twice for every withdrawal term. Which number should be larger, $P_1$ or $P_2$? Justify your response.

(c) Suppose a \$5000 annual scholarship is to be established. The Office of Advancement figures it can get an APR of 3% compounding semi-annually on any funds invested to support the scholarship. Determine the amount which must be invested today to support the fund.

2-58. In this exercise, we'll determine how to modify the annuity formula to account for discrepencies in the compounding and deposit periods. In each case, suppose a principal $R$ is invested at regular intervals, compounding at a rate $R\%$ in each period, over $N$ periods.

(a) Suppose your deposits occur three times more frequently than your compounding period (for example, if you invest monthly but compound quarterly). Write out a sum which describes the present value of this investment, then use the formula for a finite geometric series to reduce this to a formula without sums.

(b) Generalize part (a) by considering the situation where you deposit $k$-times more frequently than your compounding period.

(c) Suppose your componding period occurs three times more frequently than your deposits (for example, if you invest quarterly but compound monthly), and that $N$ is divisible by 3. Write out a sum which describes the present value of this investment, then use the formula for a finite geometric series to reduce this to a formula without sums.

(d) Generalize part (a) by considering the situation where you compound $k$-times more frequently than your deposit period. Again, you may assume that $N$ is divisible by $k$.

2-59. With an ordinary annuity (compounding at the end of the each term), if the compounding and deposit periods coincide, we know that the future value of that account with payments $R$ compounding at a rate $r$ after $N$ terms is

$$S = \sum_{k=0}^{N} R(1+r)^k = R\frac{(1+r)^N - 1}{r}.$$

(a) How does this change if payments are made at the beginning of each term, instead of the end?

(b) Evaluate the new geometric series.

2-60. Today is January 1st. You invest a principle $R$ at the end of each month. The annuity compounds at a rate $r$ at the end of odd months, and a rate $s$ at the end of odd months.

(a) Express the future value of the account after $2N$ months as a sum.

(b) Evaluate that sum using the expression for a finite geometric series.

2-61. As a bank employee, you've been asked to determine the pricing of an annuity for a client. The client would like to receive lump sum payments of $\$50,000$ at the *beginning* of each year for the next 10 years. The first payment would be made immediately upon purchasing the annuity, so that there are exactly 10 payments.

(a) The bank wants to make a profit off selling this annuity. If the bank calculates it can make an annualized return of $r$ off the initial lump sum payment, should the bank calculate the cost of the annuity at a rate higher or lower than $r$?

(b) If the bank calculates annuities at an APR of 3% compounding at the end of each *month*, what should you charge the client?

(c) To make your life easier in the future, you decide to calculate the charge for this annuity if the yearly payment is $R$ dollars over $N$ years ($N$ total payments). What is the charge for such an annuity?

(d) The original client changed their mind, and wishes to purchase a perpetuity instead. What should you charge them?

# 3 Linear Algebra

Linear Algebra is the study of systems of equations, and as such is a critical field of mathematics with wide-sweeping applications. For example, things like portfolio optimization, earthquake detection, computer generated graphics, and even Google itself all depend upon linear algebra. Linear Algebra appears almost any time more than a single variable is involved, and given that most of real life requires more than one variable (we exist in *three* dimensions for example), it shows up a lot.

You've likely seen examples of linear systems before, such as the system below:

$$\begin{aligned} 2x - 3y &= -7 \\ -x + 2y &= \phantom{-}5. \end{aligned}$$

Here there are two equations and two variables. The system is sufficiently simple that we might even be able to guess an answer, but what if we add more equations and more unknowns? Something along the lines of

$$\begin{aligned} x + \phantom{2}y + \phantom{3}z - 8w &= 4 \\ x + 2y + 3z \phantom{{}+ w} &= 9 \\ 2x + 3y + \phantom{3}z + \phantom{8}w &= 7 \end{aligned}$$

is much more difficult. You can imagine systems consisting of four equations, five equations, and it's not unusual to see systems with millions of equations and millions of unknowns. We will develop a scheme for solving these types of systems.

## 3.1 Linear Equations and Systems

A *linear equation* is any equation of the form

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = b,$$

for $c_1, c_2, \ldots, c_n, b \in \mathbb{R}$. We refer to the $c_i$ as *coefficients* of the linear equation, the $x_i$ are the *variables*, and $b$ is the *constant term*. When $n = 2$ this becomes the equation

$$c_1 x_1 + c_2 x_2 = b,$$

and the collection of $x_1$ and $x_2$ which satisfy this equation form a line in the plane. For example,

$$2x_1 - 3x_2 = -7$$

looks like the line given in Figure 3.1.

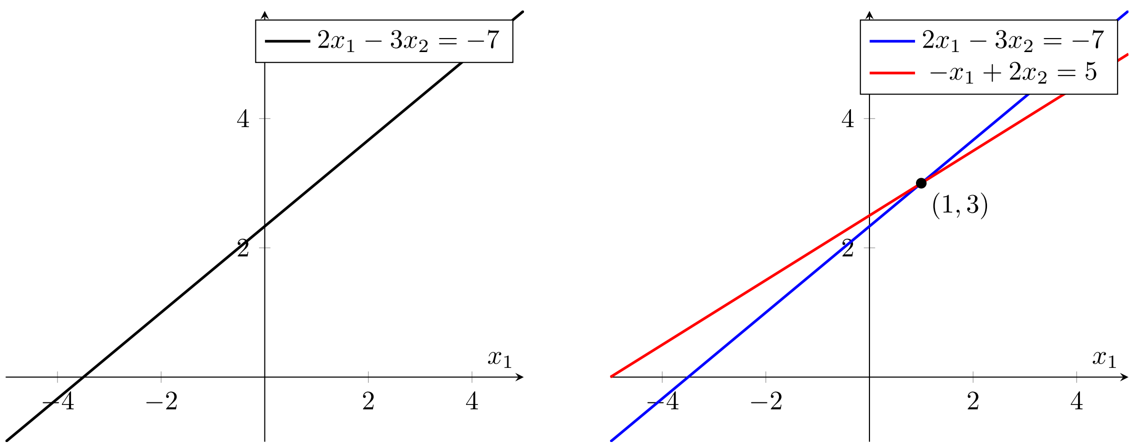



Figure 3.1: Left: The solutions to $2x_1 - 3x_2 = -7$ form a line in the plane. Right: There is a single solution to the system (3.1), which is the point where the two solution sets intersect.

A *linear system of equations* is a finite collection of linear equations,

$$\begin{array}{rcl} a_{1,1}x_1 + \; a_{1,2}x_2 + \cdots + \; c_{1,n}x_n &=& b_1 \\ a_{2,1}x_1 + \; a_{2,2}x_2 + \cdots + \; c_{2,n}x_n &=& b_2 \\ \vdots \qquad\quad \vdots \qquad\quad \vdots & & \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + c_{m,n}x_n &=& b_m \end{array}.$$

This particular system has $m$ equations in $n$ unknown, and we note that $m$ and $n$ need not be the same number. A solution to this system is any collection of $n$ numbers $s_1, s_2, \ldots s_n$ such that

$$\begin{array}{rcl} a_{1,1}s_1 + \; a_{1,2}s_2 + \cdots + \; c_{1,n}s_n &=& b_1 \\ a_{2,1}s_1 + \; a_{2,2}s_2 + \cdots + \; c_{2,n}s_n &=& b_2 \\ \vdots \qquad\quad \vdots \qquad\quad \vdots & & \\ a_{m,1}s_1 + a_{m,2}s_2 + \cdots + c_{m,n}s_n &=& b_m \end{array};$$

that is, each equation is satisfied by the $s_1, \ldots, s_n$ *simultaneously*. For example, recall the system above:

$$\begin{aligned} 2x_1 - 3x_2 &= -7 \\ -x_1 + 2x_2 &= \phantom{-}5. \end{aligned} \tag{3.1}$$

The point $(-2, 1)$ satisfies the first equation,

$$2(-2) - 3(1) = -4 - 3 = -7,$$

but does **not** satisfy the second equation, as

$$-(-2) + 2(1) = 4 + 2 = 6 \neq 5.$$

In fact, the only simultaneous solution to both equations is $(x_1, x_2) = (1, 3)$, visualized in Figure 3.1.

**Remark 3.1** Students often ask why we write $2x_1 + 3x_2 = 7$ instead of $2x + 3y = 7$. There is no difference between these two equations, since the variables $(x_1, x_2)$ or $(x, y)$ are just place-holders for particular values. We just as easily could write $2\xi + 3\zeta = 7$ or $2\spadesuit + 3\heartsuit = 7$; however, when there are a large – or indeterminate – number of variables, it is easier to simply label them as "variable 1, variable 2, variable 3, etc" than invent new symbols for each.

### 3.1.1 Number of Solutions

A natural question arises as to the number of solutions that a system can have. We can glean some insight to this question by thinking about the two variable case. For example, every linear equation $ax + by = c$ is a line in the plane (hence the name **line**ar). As long as $b \neq 0$, we can re-arrange this into our usual equation of a line

$$y = -\frac{a}{b}x + \frac{c}{b},$$

and when $b = 0$ this is a vertical line $x = c/a$.

Now let's add a second equation to the mix, giving ourselves a linear system:

$$\begin{aligned} ax + by &= c \\ fx + gy &= h \end{aligned}.$$

A solution to this linear system is any place where both lines intersect. You'll have to think geometrically, but a pair of lines in the plane will either intersect at exactly one point, at no points (they're parallel), or at infinitely many points (they overlap).

This is a good start, though in reality the situation is more complicated than above. In particular, what happens if you have three equations in two unknowns?

$$\begin{aligned} ax + by &= c \\ fx + gy &= h. \\ jx + ky &= \ell \end{aligned}$$

There are now three lines, and they will only have a solution if all three lines intersect *at the same point*; namely, all three lines could intersect, but yet the system might fail to admit a solution (Figure 3.3).

## 3.2 Vectors and Matrices

We're going to jump away from linear systems in order to introduce new tools. We'll return once we are in a better position to analyze and solve linear systems.

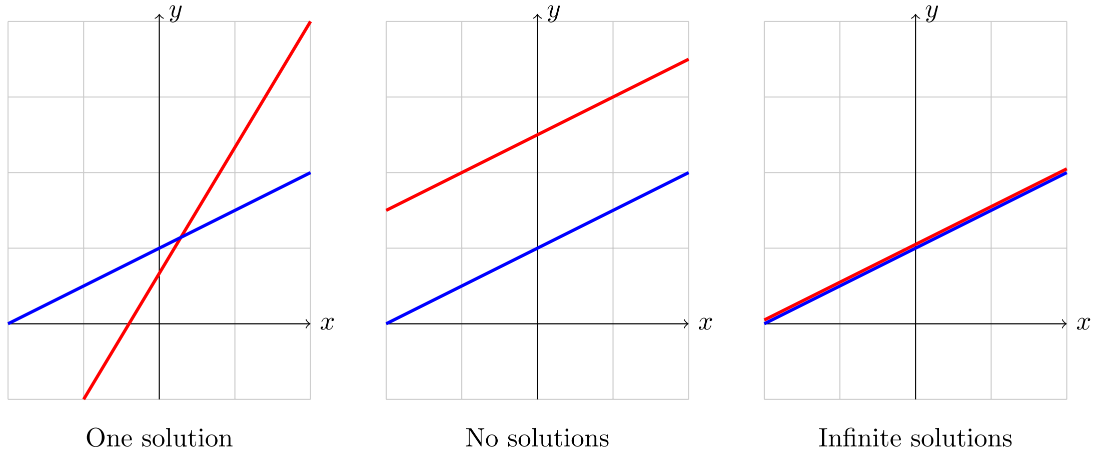



Figure 3.2: Illustrating the number of solutions for a system of two equations in two unknowns, there are can either be one solution, no solutions, or infinitely many.

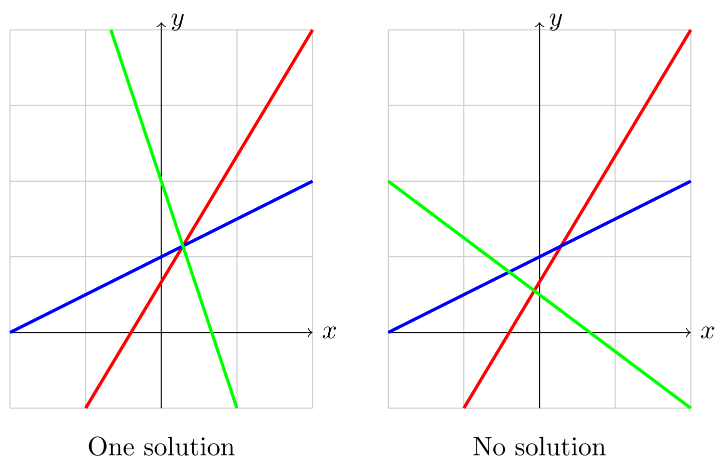



Figure 3.3: Left: Three lines in the plane intersecting at a single point means our system has a single solution. Right: Even though each pair of lines intersect, because they fail to intersect at the same point, the system fails to admit a solution.

### 3.2.1 Vectors

To discuss multiple variables, or points in multiple dimensions, we need the notion of an $n$-tuple.

**Definition 3.2**

If $n$ is a positive integer, an $n$-tuple is any collection of $n$ real numbers, written as $(a_1, a_2, \ldots, a_n)$. The set of all $n$-tuples is denoted $\mathbb{R}^n$.

So for example,

$$(-5, \pi, 1001) \in \mathbb{R}^3, \quad (0,0,1,0) \in \mathbb{R}^4, \quad \underbrace{(1,0,1,0,\cdots,1,0)}_{\text{20-times}} \in \mathbb{R}^{20}.$$

Two $n$-tuples are equal when they have precisely the same numbers, in precisely the same order. Elements in $\mathbb{R}^n$ can be thought of as either points, or arrows. For example, $(a,b) \in \mathbb{R}^2$ is either the point whose coordinates are $(a,b)$, or the arrow pointing from the origin $(0,0)$ to $(a,b)$. This is illustrated in Figure 3.4.

Thinking of $n$-tuples as arrows, we can add them together in a pointwise fashion, or multiply them by real numbers:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \qquad 2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

Multiplying by a real number – as in the latter example – is called *scalar multiplication*. Geometrically, adding two vectors is equivalent to placing the two vectors tip-to-tail and taking the new arrow that they form, while scalar multiplication amounts to scaling and reflecting a vector (see Figure 3.4). When we allow ourselves these properties of addition and scalar multiplication, we refer to $n$-tuples as *vectors*. Vectors in these notes will be denoted by a bold-face font, such as $\mathbf{b}$.

**Remark 3.3** The distinction between vectors and $n$-tuples may be confusing, and for the most part I would encourage you to not worry about their difference. An $n$-tuple is a collection of numbers, while a vector is a collection of numbers *that can be added to one another and multiplied by a scalar*. It's like asking the difference between a robot and a statue: Both might have metal shells, but when you add functionality like gears to make a statue move, it becomes something different.

### 3.2.2 Matrices

An $m \times n$ matrix is a collection of $mn$ numbers, arranged into $m$ rows and $n$ columns. For example,

$$\begin{bmatrix} 1/2 & -\pi & 4 \\ 0 & 0 & \sqrt{2} \end{bmatrix}$$

is a $2 \times 3$ matrix of real numbers.

We generally denote a matrix by a capital letter, for example $A$. We denote the $(i,j)$-element ($i$th row, $j$th column) of $A$ as $A_{ij}$ and write $[A_{ij}]$ to refer to the matrix made up of these entries.

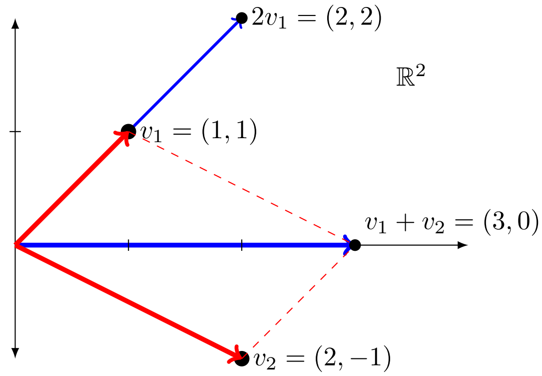



Figure 3.4: One may think of a vector as either representing a point in the plane (represented by the black dots) or as direction with magnitude (represented by the red arrows). The blue arrows correspond to the sum $v_1 + v_2$ and the scalar multiple $2v_1$. Notice that both are computed pointwise.

For example, if

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ -2 & -4 & -6 & -8 \end{bmatrix}$$

then $A_{2,3} = 7$ and $A_{3,4} = -8$.

Two matrices are equal if they are the same size and have identical elements. More precisely, if $A$ and $B$ are both $m \times n$ matrices, then $A = B$ if and only if $A_{ij} = B_{ij}$ for every $1 \leq i \leq m, 1 \leq j \leq n$. We can add two matrices of the same size by saying that $(A + B)_{ij} = A_{ij} + B_{ij}$. For example, if

$$A = \begin{bmatrix} -1 & 4 & 2 \\ 0 & -3 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 4 & 6 \\ -2 & 4 & 3 \end{bmatrix},$$

then

$$A + B = \begin{bmatrix} -1 & 4 & 2 \\ 0 & -3 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 4 & 6 \\ -2 & 4 & 3 \end{bmatrix} = \begin{bmatrix} 2 + (-1) & 4 + 4 & 2 + 6 \\ 0 + (-2) & -3 + 4 & 0 + 3 \end{bmatrix} = \begin{bmatrix} 1 & 8 & 8 \\ -2 & 1 & 3 \end{bmatrix}.$$

We can perform an operation called *scalar multiplication* by taking $c \in \mathbb{R}$ and defining $cA$ to be $(cA)_{ij} = cA_{ij}$. For example, if $c = 3$ and $A$ is as above, then

$$3A = 3 \begin{bmatrix} -1 & 4 & 2 \\ 0 & -3 & 0 \end{bmatrix} = \begin{bmatrix} -3 & 12 & 6 \\ 0 & -6 & 0 \end{bmatrix}.$$

With the ability to add and apply scalar multiplication, we note that vectors are just special cases of matrices; namely, a vector $\mathbf{v} \in \mathbb{R}^n$ is just a $1 \times n$ matrix.

**Theorem 3.4**

If $A, B, C$ are $m \times n$ matrices, with $r, s, t \in \mathbb{R}$ then

1. $A + B = B + A$
2. $A + (B + C) = (A + B) + C$
3. $0 + A = A$ (where 0 is the 0-matrix)
4. $A + (-A) = 0$ where $-A = -1A$.
5. $r(A + B) = rA + rB$
6. $(r + s)A = rA + sA$
7. $(rs)A = r(sA)$

**Example 3.5**

Suppose that

$$A = \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} x+y & 0 \\ -4 & x-y \end{bmatrix},$$

satisfy $4A + 2B = 0$. Find $x, y$.

*Solution.* By definition:

$$4A + 2B = \begin{bmatrix} -4 & 0 \\ 8 & 12 \end{bmatrix} + \begin{bmatrix} 2x+2y & 0 \\ -8 & 2x-2y \end{bmatrix} = \begin{bmatrix} 2x+2y-4 & 0 \\ 0 & 2x-2y+12 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

which means we need $2x+2y = 4$ and $2x-2y = -12$. We can solve this linear system by introducing a matrix and row reducing:

$$\left[\begin{array}{cc|c} 2 & 2 & 4 \\ 2 & -2 & -12 \end{array}\right] \xrightarrow{(-1)\mathrm{R}_1+\mathrm{R}_2\to\mathrm{R}_2} \left[\begin{array}{cc|c} 2 & 2 & 4 \\ 0 & -4 & -16 \end{array}\right] \xrightarrow[(-1/4)\mathrm{R}_2\to\mathrm{R}_2]{(1/2)\mathrm{R}_1\to\mathrm{R}_1} \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & 4 \end{array}\right]$$
$$\xrightarrow{(-1)\mathrm{R}_2+\mathrm{R}_1\to\mathrm{R}_1} \left[\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 4 \end{array}\right],$$

so $x = -2$ and $y = 4$. ■

### 3.2.3 Linear Combinations and Matrix Representations

We can now write linear systems in the language of matrices and vectors.

**Definition 3.6**

Given column vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in \mathbb{R}^n$, and scalars $c_1, c_2, \ldots, c_k$, we call anything of the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

a *linear combination* of the vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$.

Our interest lies in whether one vector can be written as a linear combination of other vectors.

For example, one might ask whether

$$\begin{bmatrix} -6 \\ 3 \\ -7 \end{bmatrix} \quad \text{can be written as a linear combination of} \quad \left\{ \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} \right\}.$$

How does this relate to linear systems? Asking that one vector can be written as a linear combination of other others is equivalent to asking if there are constants $x_1, x_2, x_3$ such that

$$\begin{bmatrix} -6 \\ 3 \\ -7 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}.$$

Using scalar multiplication and vector addition to simplify the right hand side gives

$$\begin{bmatrix} -6 \\ 3 \\ -7 \end{bmatrix} = \begin{bmatrix} x_1 \\ 3x_1 \\ 4x_1 \end{bmatrix} + \begin{bmatrix} -x_2 \\ 0 \\ 2x_2 \end{bmatrix} + \begin{bmatrix} x_3 \\ x_3 \\ 3x_3 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 + x_3 \\ 3x_1 + x_3 \\ 4x_1 + 2x_2 + 3x_3 \end{bmatrix}.$$

Since two vectors are equal precisely when they have the same numbers, this means we get a linear system of equations:

$$\begin{aligned} x_1 - \ x_2 + \ x_3 &= -6 \\ 3x_1 - \qquad + \ x_3 &= 3 \\ 4x_1 + 2x_2 + 3x_3 &= -7 \end{aligned}$$

This process can also be reversed. Consider the linear system

$$\begin{aligned} x_1 + 2x_2 - 4x_3 &= 10 \\ 2x_1 - \ x_2 + 2x_3 &= \ 5. \\ x_1 + \ x_2 - 2x_3 &= \ 7 \end{aligned}$$

Define column vectors whose elements are the coefficients of each $x_i$

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} -4 \\ 2 \\ -2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 10 \\ 5 \\ 7 \end{bmatrix}.$$

Thinking of the $x_i$ as scalars, our linear system above is equivalent to

$$\begin{aligned} \mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \mathbf{a}_3 x_3 = \mathbf{b} \quad &\Leftrightarrow \quad \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} x_1 + \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} x_2 + \begin{bmatrix} -4 \\ 2 \\ -2 \end{bmatrix} x_3 = \begin{bmatrix} 10 \\ 5 \\ 7 \end{bmatrix} \\ &\Leftrightarrow \quad \begin{bmatrix} x_1 + 2x_2 - 4x_3 \\ 2x_1 - x_2 + 2x_3 \\ x_1 + x_2 - 2x_3 \end{bmatrix} = \begin{bmatrix} 10 \\ 5 \\ 7 \end{bmatrix}. \end{aligned}$$

Alternatively, matrices can be used to encode information about linear systems. Given the system

$$\begin{matrix} a_{1,1}s_1 + & a_{1,2}s_2 + \cdots + & c_{1,n}s_n = & b_1 \\ a_{2,1}s_1 + & a_{2,2}s_2 + \cdots + & c_{2,n}s_n = & b_2 \\ \vdots & \vdots & \vdots & \\ a_{m,1}s_1 + & a_{m,2}s_2 + \cdots + & c_{m,n}s_n = & b_m \end{matrix},$$

we encode this information in an (augmented) $m \times (n+1)$ matrix whose entries are the coefficients and constant terms:

$$\left[\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} & b_m \end{array}\right].$$

For example,

$$\begin{aligned} 2x - 3y &= -7 \\ -x + 2y &= \phantom{-}5. \end{aligned} \qquad \left[\begin{array}{rr|r} 2 & -3 & -7 \\ -1 & 2 & 5 \end{array}\right]$$

$$\begin{aligned} x + \phantom{2}y + \phantom{3}z &= 4 \\ x + 2y + 3z &= 9 \\ 2x + 3y + \phantom{3}z &= 7 \end{aligned} \qquad \left[\begin{array}{rrr|r} 1 & 1 & 1 & 4 \\ 1 & 2 & 3 & 9 \\ 2 & 3 & 1 & 7 \end{array}\right] .$$

$$\begin{aligned} x_1 + 2x_3 - 4x_4 &= 1 \\ x_2 - 3x_3 + \phantom{4}x_4 &= 0 \end{aligned} \qquad \left[\begin{array}{rrrr|r} 1 & 0 & 2 & -4 & 1 \\ 0 & 1 & -3 & 1 & 0 \end{array}\right]$$

## 3.3 Solving Linear Systems

Now let's think about what operations we can do to our system of equations while preserving the solutions, and see how those operations translate to the matrix picture.

1. **We can interchange any two equations.** Certainly it does not matter whether we solve the system

$$\begin{aligned} 2x - 3y &= -7 \\ -x + 2y &= \phantom{-}5 \end{aligned} \quad \text{or} \quad \begin{aligned} -x + 2y &= \phantom{-}5 \\ 2x - 3y &= -7\text{'} \end{aligned}$$

   so we can interchange the rows of a matrix,

$$\left[\begin{array}{rr|r} 2 & -3 & -7 \\ -1 & 2 & 5 \end{array}\right] \xrightarrow{\mathrm{R}_1 \leftrightarrow \mathrm{R}_2} \left[\begin{array}{rr|r} -1 & 2 & 5 \\ 2 & -3 & -7 \end{array}\right].$$

2. **We can multiply a row by a non-zero number.** For example, if $s_1, s_2$ satisfy

$$2s_1 - 3s_2 = -7,$$

   then multiplying everything by 5 gives

$$10s_1 - 15s_2 = -35.$$

   So long as the coefficients *and* the constant term are both multiplied by the same constant, $(s_1, s_2)$ is still a solution.

$$\left[\begin{array}{rr|r} 2 & -3 & -7 \\ -1 & 2 & 5 \end{array}\right] \xrightarrow{5\mathrm{R}_1 \to \mathrm{R}_1} \left[\begin{array}{rr|r} 10 & -15 & -35 \\ -1 & 2 & 5 \end{array}\right]$$

3. **We can add a multiple of one row to another.** For example, if $(s_1, s_2)$ is a solution to

$$\begin{aligned} 2s_1 - 3s_2 &= -7 \\ -s_1 + 2s_2 &= \phantom{-}5 \end{aligned}$$'

then taking 3 times the first row and adding it to the second gives

$$\begin{array}{rrrrrr} & 3( & 2s_1 & - & 3s_2 & = & -7) \\ + & & -s_1 & + & 2s_2 & = & 5 \\ \hline & & 5s_1 & - & 7s_2 & = & -16. \end{array}$$

At the matrix level, we get

$$\left[\begin{array}{rr|r} 2 & -3 & -7 \\ -1 & 2 & 5 \end{array}\right] \xrightarrow{3R_1+R_2 \to R_2} \left[\begin{array}{rr|r} 2 & -3 & -7 \\ 5 & -7 & -16 \end{array}\right]$$

These are called *elementary row operations* (EROs).

How does this help us solve linear systems? At the moment, these matrices represent a notation for convenient bookkeeping, so let's compare this to how we would normally solve this system. Take the system

$$\begin{aligned} 2x_1 - 3x_2 &= -7 \\ -x_1 + 2x_2 &= \phantom{-}5 \end{aligned}.$$

We add twice the second row to the first to get

$$\begin{array}{rrrrrrr} & & 2x_1 & - & 3x_2 & = & -7 \\ + & 2( & -x_1 & + & 2x_2 & = & 5) \\ \hline & & & & x_2 & = & 3 \end{array}$$

from which we conclude $x_2 = 3$. The point of this particular operation was to "eliminate variables;" namely, through an adept combination of scalar multiplication and addition, we were able to eliminate $x_1$ from the second equation. Now knowing that $x_2 = 3$, we can substitute this back into the equation $2x_1 - 3x_2 = -7$ to get

$$2x_1 - 3(3) = -7 \qquad \Rightarrow \qquad 2x_1 = 2 \qquad \Rightarrow \qquad x_1 = 1$$

and we get the solution $(x_1, x_2) = (1, 3)$.

**Definition 3.7**

An $m \times n$ matrix is said to be in *row-echelon form* (REF) if

1. Any row consisting of entirely zeros appears at the bottom of the matrix,
2. The first non-zero entry of any row is a 1, called the *leading* 1,
3. Each leading 1 occurs to the right of any leading 1 which occurs above it.

Moreover, a matrix is said to be in *reduced* row-echelon form (RREF) if it is in REF, and moreover

4. Each leading one is the only non-zero element in its column.

For example,

$$\begin{bmatrix} \mathbf{1} & 2 & 3 & 4 & 5 \\ 0 & 0 & \mathbf{1} & 6 & 7 \\ 0 & 0 & 0 & \mathbf{1} & 8 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{is in REF, while} \quad \begin{bmatrix} \mathbf{1} & 2 & 0 & 0 & 5 \\ 0 & 0 & \mathbf{1} & 0 & 7 \\ 0 & 0 & 0 & \mathbf{1} & 8 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{is in RREF.} \tag{3.2}$$

The presence of a 0 in a matrix means its corresponding variable does not appear in the corresponding linear system. To see this, note the following matrix is in REF,

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \text{and corresponds to the linear system} \quad \begin{aligned} x_1 + 2x_2 + 3x_3 &= \phantom{-}4 \\ x_3 &= -1 \end{aligned}$$

Our goal then is to use the elementary row operations to turn as many of the elements of a matrix into zeros as possible. Row-echelon form represents a state in which the system has been reduced to make it amenable to back substitution, while reduced row echelon form is the maximally reduced form of the matrix.

### 3.3.1 Gaussian Elimination

The process of turning a matrix into row-echelon form – and eventually into reduced row-echelon form – is called *Gaussian elimination*. Using the elementary row operations, we progressively reduce our matrix so that the lower left triangle consists of as many zeros as possible. If we desire reduced row echelon form, we perform a similar series of steps to reduce the upper right corner to as many zeroes as possible.

It's important to note that there are many different ways of performing the EROs, with some more clever than others. It's important to practice a great deal to get a feel for how the algorithm is performed.

**Example 3.8**

Perform elementary row operations on the augmented matrix

$$\left[\begin{array}{cc|c} 2 & -3 & -7 \\ -1 & 2 & 5 \end{array}\right]$$

to turn it into row-echelon form.

*Solution.* It's easier to work with the 1 in the second row, so we'll switch the first and second rows to make our lives easier. We'll then use our elementary row operations to turn the $(2,1)$ element into a zero.

$$\left[\begin{array}{cc|c} 2 & -3 & -7 \\ -1 & 2 & 5 \end{array}\right] \xrightarrow{\mathrm{R}_1 \leftrightarrow \mathrm{R}_2} \left[\begin{array}{cc|c} -1 & 2 & 5 \\ 2 & -3 & -7 \end{array}\right] \xrightarrow{(-1)\mathrm{R}_1 \to \mathrm{R}_1} \left[\begin{array}{cc|c} 1 & -2 & -5 \\ 2 & -3 & -7 \end{array}\right]$$
$$\xrightarrow{(-2)\mathrm{R}_1 + \mathrm{R}_2 \to \mathrm{R}_2} \left[\begin{array}{cc|c} 1 & -2 & -5 \\ 0 & 1 & 3 \end{array}\right].$$

Converting back into the corresponding linear system gives

$$\begin{aligned} x_1 - 2x_2 &= -5 \\ x_2 &= \ \ 3 \end{aligned}.$$

Knowing that $x_2 = 3$ we can solve for $x_1 = -5 + 2x_2 = 1$, which is the same solution we got earlier.

We can take this one step further, and turn the matrix into its reduced row echelon form. We'll start from the bottom and work our way upwards.

$$\left[\begin{array}{rr|r} 2 & -3 & -7 \\ -1 & 2 & 5 \end{array}\right] \xrightarrow{\text{REF}} \left[\begin{array}{rr|r} 1 & -2 & -5 \\ 0 & 1 & 3 \end{array}\right] \xrightarrow{2R_2+R_1 \to R_1} \left[\begin{array}{rr|r} 1 & 0 & 1 \\ 0 & 1 & 3 \end{array}\right].$$

The corresponding linear system is just the solution $(x_1, x_2) = (1, 3)$. ■

**Example 3.9**

Use Gaussian elimination to convert the augmented matrix

$$\left[\begin{array}{rrr|r} 1 & 1 & 1 & 4 \\ 1 & 2 & 3 & 9 \\ 2 & 3 & 1 & 7 \end{array}\right]$$

into row-echelon form. Use backwards substitution to solve the system. By turning the matrix into reduced row-echelon form, confirm your answer.

*Solution.* Starting at $(1, 1)$ entry, we will use the elementary row operations to turn the $(2, 1)$ and $(3, 1)$ elements to 0. After this, we will move to the second column and perform a similar operation. Applying Gaussian elimination, we get

$$\left[\begin{array}{rrr|r} 1 & 1 & 1 & 4 \\ 1 & 2 & 3 & 9 \\ 2 & 3 & 1 & 7 \end{array}\right] \xrightarrow[(-2)R_1+R_3 \to R_3]{(-1)R_1+R_2 \to R_2} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 4 \\ 0 & 1 & 2 & 5 \\ 0 & 1 & -1 & -1 \end{array}\right] \xrightarrow{(-1)R_2+R_3 \to R_3} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 4 \\ 0 & 1 & 2 & 5 \\ 0 & 0 & -3 & -6 \end{array}\right]$$

$$\xrightarrow{(-1/3)R_3 \to R_3} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 4 \\ 0 & 1 & 2 & 5 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

The corresponding linear system is

$$\begin{aligned} x_1 + x_2 + \ \ x_3 &= 4 \\ x_2 + 2x_3 &= 5. \\ x_3 &= 2 \end{aligned}$$

Setting $x_3 = 3$ and substituting into the second equation gives $x_2 = 5 - 2x_3 = 1$. Substituting both values into the first equation gives

$$x_1 = 4 - x_2 - x_3 = 4 - (1) - (2) = 1$$

so the solution is $(x_1, x_2, x_3) = (1, 1, 2)$. You can check the answer by substituting into the original linear system of equations.

To turn this into reduced row-echelon form, we start at the $(3,3)$ entry of our row-echelon form and work upwards:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 0 & 1 & 2 & 5 \\ 0 & 0 & 1 & 2 \end{array}\right] \xrightarrow[(-1)R_1+R_1\to R_1]{(-2)R_3+R_2\to R_2} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right] \xrightarrow{(-1)R_2+R_1\to R_1} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

which gives us the same solution above, $(x_1, x_2, x_3) = (1, 1, 2)$. ■

**Example 3.10**

A toy producer manufactures three different products $A, B$, and $C$. The cost to produce each unit is $[\$3, \$6, \$8]$ respectively, while the profit from each sale is $[\$5, \$3, \$10]$. You have a $\$60,000$ budget, need to produce 8000 units total, and are aiming for a quarterly profit of $\$40,000$. Fixed costs run at $\$11,000$. Determine how much of each toy you should produce.

*Solution.* Let $x, y$, and $z$ denote the quantity of $A$,$B$, and $C$ to produce. We have three equations: the total cost, the desired profit, and the total number of units to produce. The total cost is

$$\begin{aligned} \text{Total Cost} &= (\text{Fixed Cost}) + (\text{Cost Per Unit}) \\ 60000 &= 11000 + 3x + 6y + 8z \\ 49000 &= 3x + 6y + 8z. \end{aligned}$$

The total profit is the sum of the profits from each unit, so $40000 = 5x + 3y + 10z$, while the total number of units produced is $x + y + z = 8000$. Putting this together gives us the linear system

$$\begin{aligned} x + \phantom{6}y + \phantom{10}z &= 8000 \\ 3x + 6y + \phantom{1}8z &= 49000. \\ 5x + 3y + 10z &= 40000 \end{aligned}$$

Converting this to matrix form and putting into reduced row echelon, we get

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 8000 \\ 3 & 6 & 8 & 49000 \\ 5 & 3 & 10 & 40000 \end{array}\right] \xrightarrow[-5R_1+R_3\to R_3]{-3R_1+R_2\to R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 8000 \\ 0 & 3 & 5 & 25000 \\ 0 & -2 & 5 & 0 \end{array}\right] \xrightarrow[(3/25)R_3\to R_3]{(2/3)R_2+R_3\to R_3} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 8000 \\ 0 & 3 & 5 & 25000 \\ 0 & 0 & 1 & 2000 \end{array}\right]$$

$$\xrightarrow[-1R_1+R_1\to R_1]{-5R_1+R_2\to R_2} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 6000 \\ 0 & 3 & 0 & 15000 \\ 0 & 0 & 1 & 2000 \end{array}\right] \xrightarrow[-1R_2+R_1\to R_1]{(1/3)R_2\to R_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1000 \\ 0 & 1 & 0 & 5000 \\ 0 & 0 & 1 & 2000 \end{array}\right].$$

Hence the company should make 1000 units of $A$, 5000 units of $B$, and 2000 units of $C$. ■

**Example 3.11**

Determine whether $\mathbf{v}$ can be written as a linear combination of $\mathbf{x}, \mathbf{y}, \mathbf{z}$, where

$$\mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 5 \\ 3 \\ 4 \end{bmatrix}.$$

*Solution.* If a solution exists, there are $x_1, x_2, x_3$ such that $\mathbf{v} = x_1\mathbf{x} + x_2\mathbf{y} + x_3\mathbf{z}$, which is the same as solving the linear system

$$\begin{aligned} 2x_1 + x_2 + \phantom{2}x_3 &= 5 \\ x_1 \phantom{+ x_2} + \phantom{2}x_3 &= 3 \\ -x_1 + x_2 + 2x_3 &= 4 \end{aligned} \qquad \text{with matrix} \qquad \left[\begin{array}{rrr|r} 2 & 1 & 1 & 5 \\ 1 & 0 & 1 & 3 \\ -1 & 1 & 2 & 4 \end{array}\right]$$

Applying Guassian elimination gives

$$\left[\begin{array}{rrr|r} 2 & 1 & 1 & 5 \\ 1 & 0 & 1 & 3 \\ -1 & 1 & 2 & 4 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{rrr|r} 1 & 0 & 1 & 3 \\ 2 & 1 & 1 & 5 \\ -1 & 1 & 2 & 4 \end{array}\right] \xrightarrow[R_1 + R_3 \to R_3]{(-2)R_1 + R_2 \to R_2} \left[\begin{array}{rrr|r} 1 & 0 & 1 & 3 \\ 0 & 1 & -1 & -1 \\ 0 & 1 & 3 & 7 \end{array}\right]$$

$$\xrightarrow{(-1)R_2 + R_3 \to R_3} \left[\begin{array}{rrr|r} 1 & 0 & 1 & 3 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 4 & 8 \end{array}\right] \xrightarrow{(1/4)R_3 \to R_3} \left[\begin{array}{rrr|r} 1 & 0 & 1 & 3 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

$$\xrightarrow[R_3 + R_2 \to R_2]{(-1)R_3 + R_1 \to R_1} \left[\begin{array}{rrr|r} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right].$$

This does indeed have a solution, showing that $\mathbf{x} + \mathbf{y} + 2\mathbf{z} = \mathbf{v}$, so $\mathbf{v}$ is a linear combination of $\mathbf{x}, \mathbf{y}$, and $\mathbf{z}$. ■

**Example 3.12**

A mutual fund consists of three funds: An aggressive (A), moderate (M), and low risk (L) portfolio. These three portfolios diversify according to three index funds: Canadian Bonds (B), Canadian Equity (C), and International Equity (I). The proportion of each is given by the following table:

| | A | M | L |
|---|---|---|---|
| B | 1 | 3 | 8 |
| C | 3 | 3 | 1 |
| I | 6 | 3 | 1 |

You decide to create a custom mix for your portfolio, and wish it to consist of 610 units of $B$, 500 units of $C$, and 860 units of $I$. How much of each fund should you buy?

*Solution.* Let $[x, y, z]$ be the amount of $A$, $M$, and $L$ bought respectively, with $\mathbf{b} = [610, 500, 860]$. We want to know if there is a linear combination of $A, M, L$ that gives $\mathbf{b}$, equivalent to solving the linear system $xA + yM + zL = \mathbf{b}$, or

$$\begin{aligned} x + 3y + 8z &= 610 \\ 3x + 3y + \phantom{8}z &= 500. \\ 6x + 3y + \phantom{8}z &= 860 \end{aligned}$$

Setting up the augmented matrix and row reducing gives

$$\left[\begin{array}{ccc|c} 1 & 3 & 8 & 610 \\ 3 & 3 & 1 & 500 \\ 6 & 3 & 1 & 860 \end{array}\right] \xrightarrow[-6R_1+R_3\to R_3]{-3R_1+R_2\to R_2} \left[\begin{array}{ccc|c} 1 & 3 & 8 & 610 \\ 0 & -6 & -23 & -1330 \\ 0 & -15 & -47 & -2800 \end{array}\right] \xrightarrow[-5R_2\to R_2]{-2R_3\to R_3} \left[\begin{array}{ccc|c} 1 & 3 & 8 & 610 \\ 0 & 30 & 115 & 6650 \\ 0 & 30 & 94 & 5600 \end{array}\right]$$

$$\xrightarrow[(-1/21)R_3\to R_3]{(-1)R_2+R_3\to R_3} \left[\begin{array}{ccc|c} 1 & 3 & 8 & 610 \\ 0 & 30 & 115 & 6650 \\ 0 & 0 & 1 & 50 \end{array}\right] \xrightarrow[-8R_3+R_1\to R_1]{-115R_3+R_2\to R_2} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 210 \\ 0 & 30 & 0 & 900 \\ 0 & 0 & 1 & 50 \end{array}\right]$$

$$\xrightarrow[-3R_2+R_1\to R_1]{(1/30)R_2\to R_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 120 \\ 0 & 1 & 0 & 30 \\ 0 & 0 & 1 & 50 \end{array}\right].$$

Hence you should by 120 units of $A$, 30 units of $M$, and 50 units of $L$. ■

### 3.3.2 Systems with No Solution

Examples 3.9, 3.10, and 3.11 all admitted unique solutions. We discussed in Section 3.1.1 that there are two other notable cases: Systems which admit no solutions, and those which admit infinitely many.

Systems with no solutions are relatively straightforward to identify. After using Gaussian elimination to convert a matrix into row-echelon form, if a row of the form $\left[\begin{array}{cccc|c} 0 & 0 & \cdots & 0 & 1 \end{array}\right]$ appears, the system has no solution. To see why, recognize that this row corresponds to the linear system $0x_1 + 0x_2 + \cdots 0x_k = 1$, or equivalently $0 = 1$. As this is impossible, the system has no solutions.

**Example 3.13**

Find the solutions to the linear system

$$\begin{aligned} x_1 - \phantom{2}x_2 + 3x_3 &= -4 \\ 2x_2 \phantom{- 2x_2} + 4x_3 &= \phantom{-}0. \\ -4x_2 - 2x_2 - 6x_3 &= \phantom{-}8 \end{aligned}$$

*Solution.* Coding this system as a matrix and converting to row echelon form gives:

$$\left[\begin{array}{ccc|c} 1 & -1 & 3 & -4 \\ 2 & 0 & 4 & 0 \\ -4 & -2 & -6 & 8 \end{array}\right] \xrightarrow[4R_1+R_3\to R_3]{(-2)R_1+R_2\to R_2} \left[\begin{array}{ccc|c} 1 & -1 & 3 & -4 \\ 0 & 2 & -2 & 8 \\ 0 & -6 & 6 & -8 \end{array}\right] \xrightarrow[3R_2+R_3\to R_3]{(1/2)R_2\to R_2} \left[\begin{array}{ccc|c} 1 & -1 & 3 & -4 \\ 0 & 1 & -1 & 4 \\ 0 & 0 & 0 & 16 \end{array}\right]$$

The final row is equivalent to the statement $0 = 16$, which is nonsense. Hence we conclude that this system admits no solutions. ■

### 3.3.3 Systems with Infinitely Many Solutions

The instance of a system which admits infinitely many solutions is a bit trickier. For example, we know that a system with two-variables might have an infinite solution set which forms a line in $\mathbb{R}^2$, and we know that we can write lines in $ax+by=c$. The problem arises when we want to write lines in three or higher dimensions. Your first guess might be to write something like $ax+by+cz=d$, but this is actually a *plane* in $\mathbb{R}^3$, not a line. In fact, to describe a line like this, we need to describe it as the intersection of two planes, which seems like way too much work. This problem continues in higher dimensions as well: What if our solution set is a "line" or a "plane" in $\mathbb{R}^4$, etc?

The way to do this is to use a small collection of vectors to describe a set of fundamental directions that one can move within the solution set. For example, in $\mathbb{R}^2$, the line which passes through the point $P=(a,b)$ in the direction $\mathbf{v}=(u,v)$ can be described the *parameterized equation*

$$L(t) = P + t\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix} + t\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} a+tu \\ b+tv \end{bmatrix}.$$

This is different way of thinking about lines than our usual $y=mx+b$ approach. The equation $y=mx+b$ describes a relationship between the $x$ and $y$ variables, so that if you know one of the variables you can find the other. The parameterized equation makes each of the $x$ and $y$ coordinates a function of $t$ ($x(t)=a+tu, y(t)=b+tv$), so that if you're given a value of $t$, you can just read off the $x$ and $y$ values.

For example, to write the equation $y=3x+2$ in parameterized form, we need to find a point through which it passes, and determine the direction in which it's travelling (Figure 3.5). Any point will do, but a good candidate is the point $P=(0,2)$. To determine the direction of travel, we can subtract two points on the line, say $Q=(1,5)$ and $P=(0,2)$ to get $\mathbf{v}=Q-P=(1,3)$. Thus the parameterized form of this line is

$$L(t) = P + t\mathbf{v} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} + t\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} t \\ 3t+2 \end{bmatrix}.$$

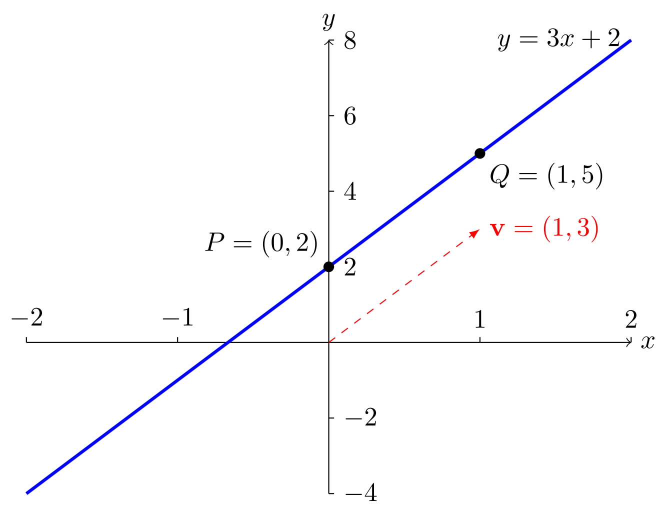



Figure 3.5: Parameterizing a line of the form $y=mx+b$.

A line is a one-parameter object, since it can be described with a single parameter ($t$). A plane is a two-parameter object, for if it passes through a point $\mathbf{p}$, then with two directional vectors $\mathbf{u}$ and $\mathbf{v}$, it is parameterized as $P(t)=\mathbf{p}+t\mathbf{v}+s\mathbf{u}$. Similarly, there are three- and four-parameter objects in higher dimensional spaces.

Our goal then is to describe infinite solutions sets in parametric form. To do this, reduce a linear system to either REF or RREF. Any column which *does not* consist of a leading one corresponds to a variable that will be made into a parameter. For example, if your linear system reduces to

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right] \qquad \text{or equivalently} \qquad \begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ x_3 &= 2 \end{aligned}$$

then the only variable without a leading one is the second column, or $x_2$. Hence we set $x_2$ to be a parameter, say $x_2 = t$. Now we solve the system with this assumption, to get

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 - 2x_2 + x_3 \\ x_2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 - 2t + 2 \\ t \\ 2 \end{bmatrix} = \begin{bmatrix} 3 - 2t \\ t \\ 2 \end{bmatrix}.$$

And indeed, we can check that this is a solution by substituting back into the equation, and hence this works for any value of $t$.

**Example 3.14**

Find the solution(s) to the linear system

$$\begin{aligned} x_1 + 2x_2 - 4x_3 &= 10 \\ 2x_1 - x_2 + 2x_3 &= 5. \\ x_1 + x_2 - 2x_3 &= 7 \end{aligned}$$

*Solution.* I'm going straight to RREF, but you are free to do backwards substitution if you like.

$$\left[\begin{array}{ccc|c} 1 & 2 & -4 & 10 \\ 2 & -1 & 2 & 5 \\ 1 & 1 & -2 & 7 \end{array}\right] \xrightarrow[(-1)R_1+R_3\to R_3]{(-2)R_1+R_2\to R_2} \left[\begin{array}{ccc|c} 1 & 2 & -4 & 10 \\ 0 & -5 & 10 & -15 \\ 0 & -1 & 2 & -3 \end{array}\right] \xrightarrow{(-1/5)R_2\to R_2} \left[\begin{array}{ccc|c} 1 & 2 & -4 & 10 \\ 0 & 1 & -2 & 3 \\ 0 & -1 & 2 & -3 \end{array}\right]$$

$$\xrightarrow{R_2+R_3\to R_3} \left[\begin{array}{ccc|c} 1 & 2 & -4 & 10 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{(-2)R_2+R_1\to R_1} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

We cannot quite read off the solutions immediately. Instead, notice that there is no leading one for the third column. This means that $x_3$ is a free parameter, say $s$. Rewriting this but solving for $x_1$ and $x_2$ gives

$$\begin{aligned} x_1 &= 4 \\ x_2 &= 3 + 2s \\ x_3 &= s, \end{aligned}$$

so the final solution is $(x_1, x_2, x_3) = (4, 3 + 2s, s)$ for any $s \in \mathbb{R}$. ∎

### 3.3.4 The Rank of a Matrix

We've seen that in describing how many solutions a system has, the number of leading ones it possesses plays an important role. For this reason, we define a special characteristic of a matrix which effectively counts its leading ones.

**Definition 3.15**

Let $A$ be a matrix. The *rank* of $A$ is the number of leading 1's in its row-echelon form.

**Example 3.16**

Determine the rank of the matrix

$$A = \begin{bmatrix} 1 & -2 & 0 & 4 \\ 3 & 1 & 1 & 0 \\ -1 & -5 & -1 & 8 \\ 3 & 8 & 2 & -12 \end{bmatrix}.$$

*Solution.* Putting this matrix into row echelon form gives

$$\begin{bmatrix} 1 & -2 & 0 & 4 \\ 3 & 1 & 1 & 0 \\ -1 & -5 & -1 & 8 \\ 3 & 8 & 2 & -12 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -2 & 0 & 4 \\ 0 & 7 & 1 & -12 \\ 0 & -7 & -1 & 12 \\ 0 & 14 & 2 & -24 \end{bmatrix} \xrightarrow[(-2)R_2+R_3\to R_3]{R_2+R_3\to R_3} \begin{bmatrix} 1 & -2 & 0 & 4 \\ 0 & 7 & 1 & -12 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

After scaling the second row, this matrix has leadings ones in the first and second columns only, and thus has rank 2. ■

This is not a good definition of rank, for several reasons which are hard to elaborate upon right now. Instead, we introduce this concept so that we can talk about the number of solutions a system can have. Notice that Examples 3.9 and 3.16 both consist of three equations in three unknowns, but the former has a unique solution while the latter had infinitely many solutions. The difference arises because of the rank, though not in an obvious way.

**Definition 3.17**

A linear system is said to be *homogeneous* if all of its constant terms are identically 0.

Homogeneous systems are special because they always have a solution; namely, the $\mathbf{0}$ vector $\mathbf{0} = (0, \ldots, 0)$ is always a solution to a homogeneous system, known as the *trivial* solution. Since we can eliminate the possibility of a homogeneous system not having any solutions, we can classify the solutions to such a system in terms of its rank.

**Theorem 3.18**

If an $m \times n$ matrix $A$ with rank $r$ describes the **coefficient matrix** of a linear homogeneous system, then

1. The system has exactly $n - r$ 'basic' solutions, one for each parameter;

2. Every solution is a unique linear combination of those basic solutions.

**Example 3.19**

Consider the linear system

$$\begin{aligned} x_1 + 2x_2 - \phantom{2}x_3 + x_4 + \phantom{3}x_5 &= 0 \\ -x_1 - 2x_2 + 2x_3 \phantom{{}+x_4} + \phantom{3}x_5 &= 0 \\ -x_1 - 2x_2 + 3x_3 + x_4 + 3x_5 &= 0 \end{aligned}$$

Determine the basic solutions and give a formula for general solutions.

*Solution.* Writing this as an augmented matrix and row-reducing, we get

$$\left[\begin{array}{rrrrr|r} 1 & 2 & -1 & 1 & 1 & 0 \\ -1 & -2 & 2 & 0 & 1 & 0 \\ -1 & -2 & 3 & 1 & 3 & 0 \end{array}\right] \xrightarrow[\mathrm{R}_1+\mathrm{R}_3\to\mathrm{R}_3]{\mathrm{R}_1+\mathrm{R}_2\to\mathrm{R}_2} \left[\begin{array}{rrrrr|r} 1 & 2 & -1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 2 & 2 & 4 & 0 \end{array}\right]$$
$$\xrightarrow{(-2)\mathrm{R}_2+\mathrm{R}_3\to\mathrm{R}_3} \left[\begin{array}{rrrrr|r} 1 & 2 & -1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Here we have rank 2 and 5 variables, so we expect there to by $5-2=3$ basic solutions corresponding to the three parameters. The variables $x_1$ and $x_3$ have the leading ones, so let $x_2 = s$, $x_4 = t$, $x_5 = u$ and write

$$\begin{aligned} x_3 &= -x_4 - 2x_5 = -t - 2u \\ x_1 &= -2x_2 + x_3 - x_4 - x_5 \\ &= -2s + (-t-2u) - t - u \\ &= -2s - 2t - 3u. \end{aligned}$$

By factoring the $s, t, u$, we can write this as a linear combination of three vectors:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2s-2t-3u \\ s \\ -t-2u \\ t \\ u \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} s + \begin{bmatrix} -2 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} t + \begin{bmatrix} -3 \\ 0 \\ -2 \\ 0 \\ 1 \end{bmatrix} u,$$

and indeed each of the three vectors above is, by itself, a solution to the homogeneous system. ■

For non-homogeneous systems (the constant terms are not all zero), solutions are surprisingly tied to the homogeneous system. Once again, if in the **augmented matrix** you see a row of the form $\left[\begin{array}{cccc|c} 0 & 0 & \cdots & 0 & 1 \end{array}\right]$, you know there are no solutions, but if the system has solutions, the following theorem is true.

**Theorem 3.20**

Consider a non-homogeneous system of equations. The vector $\mathbf{x}$ is a solution to this system if and only if $\mathbf{x} = \mathbf{x}_h + \mathbf{x}_p$, where $\mathbf{x}_h$ is a solution to corresponding homogeneous linear system (the same system but with the constants set to 0), and $\mathbf{x}_p$ is a particular solution to the linear system.

We've already seen that a homogeneous solution admits basic solutions. To employ Theorem 3.20, we employ a similar strategy, by separating out the column vectors which correspond to the parameters. Everything left over will correspond to the particular solution.

**Example 3.21**

Solve the system

$$\begin{aligned} x_1 + 2x_2 \qquad\quad - \ x_4 &= 1 \\ -2x_1 - 3x_2 + 4x_3 + 5x_4 &= 6 \\ 2x_1 + 4x_2 \qquad\quad - 2x_4 &= 2 \end{aligned}$$

and write it as a linear combination of the solution to the homogeneous system and a particular solution.

*Solution.* Writing this as a matrix and converting to RREF yields

$$\left[\begin{array}{rrrr|r} 1 & 2 & 0 & -1 & 1 \\ -2 & -3 & 4 & 5 & 6 \\ 2 & 4 & 0 & -2 & 2 \end{array}\right] \xrightarrow[(-2)\mathrm{R}_1+\mathrm{R}_3\to\mathrm{R}_3]{2\mathrm{R}_1+\mathrm{R}_2\to\mathrm{R}_2} \left[\begin{array}{rrrr|r} 1 & 2 & 0 & -1 & 1 \\ 0 & 1 & 4 & 3 & 8 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$
$$\xrightarrow{(-2)\mathrm{R}_2+\mathrm{R}_1\to\mathrm{R}_1} \left[\begin{array}{rrrr|r} 1 & 0 & -8 & -7 & -15 \\ 0 & 1 & 4 & 3 & 8 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

The leading ones occur at $x_1$ and $x_2$, so let $x_3 = s$ and $x_4 = t$ be parameter's, so that

$$\begin{aligned} x_4 &= t \\ x_3 &= s \\ x_2 &= 8 - 4x_3 - 3x_4 = 8 - 4s - 3t \\ x_1 &= -15 + 8x_3 + 7x_4 = -15 + 8s + 7t \end{aligned}$$

or written in terms of vectors, by grouping parameters

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \underbrace{\begin{bmatrix} -15 \\ 8 \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{x}_p} + \underbrace{\begin{bmatrix} 8 \\ -4 \\ 1 \\ 0 \end{bmatrix} s + \begin{bmatrix} 7 \\ -3 \\ 0 \\ 1 \end{bmatrix} t}_{\mathbf{x}_h}.$$

As indicated, $\mathbf{x}_p$ is a particular solution to the linear system, while $\mathbf{x}_h$ is the general solution to the corresponding homogeneous system. ■

## 3.4 Other Matrix Operations

The ability to treat the coefficients and constants of a linear system as their mathematical objects avails itself of some powerful theory. The next few sections expose different operations other than addition and scalar multiplication that can be performed on matrices.

### 3.4.1 The Transpose of a Matrix

Given an $m \times n$ matrix $A = [A_{ij}]$, its transpose is the $n \times m$ matrix derived by interchanging the rows and columns. We denote the transpose by $A^T$. Hence if

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}, \quad \text{then} \quad A^T = \begin{bmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{m,n} \end{bmatrix}.$$

For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

To avoid large, awkward gaps in these notes, I will sometimes use the transpose to denote column vectors, such as the $3 \times 1$ *column vector* $\mathbf{v} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^T$.

**Theorem 3.22**

If $A, B$ are $m \times n$ matrices and $c \in \mathbb{R}$, then

1. $(A^T)^T = A$,
2. $(cA)^T = cA^T$,
3. $(A + B)^T = A^T + B^T$.

**Example 3.23**

We say that a square $m \times n$ matrix $A$ is anti-symmetric if $A + A^T = 0$, where $0$ is the zero-matrix. The trace of a matrix is the sum of its diagonal terms; that is,

$$\operatorname{Tr}(A) = A_{1,1} + A_{2,2} + \cdots + A_{n,n}.$$

Show that that the trace of an anti-symmetric matrix is zero.

*Solution.* Suppose our matrix $A$ has components $A_{i,j}$. When we take the transpose, the rows and columns interchange, so that $[A^T]_{i,j} = A_{j,i}$. But notice that the diagonal elements of a square matrix are fixed under transposition: the diagonal of the original matrix is still the diagonal of the transpose. Hence

$$[A + A^T]_{i,i} = A_{i,i} + A_{i,i} = 2A_{i,i} = 0,$$

showing that $A_{i,i} = 0$. Thus the trace is

$$\operatorname{Tr}(A) = A_{1,1} + A_{2,2} + \cdots + A_{n,n} = 0 + 0 + \cdots + 0 = 0. \qquad \blacksquare$$

### 3.4.2 Matrix Multiplication

Before looking at matrix multiplication, we first consider the dot product.

**Definition 3.24**

Let $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$ be an $1 \times n$ row vector, and $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}^T$ be an $n \times 1$ column vector. The *dot product* (inner product) of $\mathbf{x}$ and $\mathbf{y}$ is

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

**Remark 3.25**

1. Strictly speaking, the dot product is always between two column vectors or two row vectors, $\mathbf{x}$ and $\mathbf{y}$. This operation of combining row and column vectors is really a very deep thing, and the fact that it is equivalent to the dot product is a theorem. Naturally, by applying the transpose we can turn these into row or column vectors, whichever we please.

2. The dot product has a nice geometric interpretation, but we cannot yet describe it until we know how to visualize column/row vectors.

**Example 3.26**

Compute the dot products of $\mathbf{x} \cdot \mathbf{y}$ and $\mathbf{y} \cdot \mathbf{z}$, where

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}.$$

*Solution.* Applying our formulas, we have

$$\begin{aligned} \mathbf{x} \cdot \mathbf{y} &= (1 \times 2) + (0 \times -5) + (1 \times 1) = 3 \\ \mathbf{y} \cdot \mathbf{z} &= (2 \times -1) + (-5 \times 0) + (1 \times 2) = 0. \end{aligned}$$ ■

Given an $n \times k$ matrix $A$ and a $k \times m$ matrix $B$, the product $AB$ is an $n \times m$ matrix, whose $(i, j)$ entry is the dot product of the $i$th row of $A$ and the $j$th column of $B$; that is,

$$(AB)_{ij} = \sum_{r=1}^{k} A_{ik} B_{kj}.$$

Alternatively, let $\mathbf{r}_i$ be the $i$th row of $A$ (of which there are $n$), and let $\mathbf{c}_j$ be the $j$th column of $B$ (of which there are $m$). Notice that both $\mathbf{r}_i$ and $\mathbf{c}_j$ have $k$-entries, so we can take their dot product, and the matrix product $AB$ is

$$AB = \begin{pmatrix} \mathbf{r}_1 \\ \hline \mathbf{r}_2 \\ \hline \vdots \\ \hline \mathbf{r}_n \end{pmatrix} \left( \begin{array}{c|c|c|c} \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_m \end{array} \right) = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{c}_1 & \mathbf{r}_1 \cdot \mathbf{c}_2 & \cdots & \mathbf{r}_1 \cdot \mathbf{c}_m \\ \mathbf{r}_2 \cdot \mathbf{c}_1 & \mathbf{r}_2 \cdot \mathbf{c}_2 & \cdots & \mathbf{r}_2 \cdot \mathbf{c}_m \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{r}_n \cdot \mathbf{c}_1 & \mathbf{r}_n \cdot \mathbf{c}_2 & \cdots & \mathbf{r}_n \cdot \mathbf{c}_m \end{bmatrix}.$$

Explicitly multiplying two $2 \times 2$ matrices $A = [A_{ij}], B = [B_{ij}]$, we get the $2 \times 2$ matrix

$$\begin{bmatrix} a_{11} & a_{12} \\ \boxed{a_{21} \quad a_{22}} \end{bmatrix} \begin{bmatrix} \boxed{\begin{matrix} b_{11} \\ b_{21} \end{matrix}} & \begin{matrix} b_{12} \\ b_{22} \end{matrix} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ \boxed{a_{21}b_{11} + a_{22}b_{21}} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}.$$

Again, I emphasize that the $i$th row and $j$th column of the product is the dot product of the $i$th row of $A$ and the $j$th column of $B$. For example, in the $2 \times 2$ case, let us look at the second row and first column. The second row of $A$ is $\begin{bmatrix} a_{21} & a_{22} \end{bmatrix}$ while the first column of $B$ is $\begin{bmatrix} b_{11} & b_{21} \end{bmatrix}^T$. Taking their dot product gives $a_{21}b_{11} + a_{22}b_{21}$ which is indeed the $(2, 1)$ entry of the product.

**Example 3.27**

Determine the matrix product $AB$ where

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 3 & -2 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 0 & -1 \\ 2 & -3 & -1 \end{bmatrix}.$$

*Solution.* The matrix $A$ has dimension $2 \times 3$ while $B$ has dimension $3 \times 3$. Their product $AB$ can therefore be computed, and will output a $2 \times 3$ matrix. Carrying out the multiplication, we get

$$AB = \begin{bmatrix} 1+0+4 & 0+0+-6 & 3+0+-2 \\ 3+0+0 & 0+0+0 & 9+2+0 \end{bmatrix} = \begin{bmatrix} 5 & -6 & 1 \\ 3 & 0 & 11 \end{bmatrix}. \qquad \blacksquare$$

A very special type of matrix is the identity matrix. If $n$ is a positive integer, we define $I_n$ to be the $n \times n$ matrix with 1's on the diagonal and zero everywhere else; that is,

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

**Theorem 3.28**

If $A$ is $m \times n$, $B$ is $n \times k$, and $C$ is $k \times \ell$, then

1. $A(BC) = (AB)C$
2. $I_m A = A I_n = A$
3. $(AB)^T = B^T A^T$

Note the interchange of order than occurs in the transpose; $(AB)^T = B^T A^T$. In fact, this must happen to ensure that the dimensions like up correctly. Since $A$ is an $m \times n$ matrix and $B$ is $n \times k$, their product $AB$ is an $m \times k$ matrix. The transpose is $k \times m$, which comes from multiplying $B^T$ with dimension $k \times n$ against $A^T$ with dimension $n \times m$.

Furthermore, matrix multiplication is distributive:

$$A(B + C) = AB + AC \quad \text{and} \quad (A + B)C = AC + BC,$$

where the dimensions of the matrices are chosen so that this makes sense.

**Example 3.29**

Show that the matrix

$$A = \begin{bmatrix} 0 & 4 \\ 2 & 2 \end{bmatrix}$$

satisfies the equation $A^2 - 2A - 8I_2 = 0$.

*Solution.* Computing $A^2$ we get

$$A^2 = \begin{bmatrix} 0 & 4 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 0 & 4 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 0+8 & 0+8 \\ 0+4 & 8+4 \end{bmatrix} = \begin{bmatrix} 8 & 8 \\ 4 & 12 \end{bmatrix}$$

so that

$$A^2 - A - 10I_2 = \begin{bmatrix} 8 & 8 \\ 4 & 12 \end{bmatrix} - 2\begin{bmatrix} 0 & 4 \\ 2 & 2 \end{bmatrix} - \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix} = \begin{bmatrix} 8-0-8 & 8-8-0 \\ 4-8-0 & 12-4+8 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

as required. ■

Something very nice happens when we multiply a matrix and a column vector. Suppose that

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}, \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

in which case the matrix product $A\mathbf{x}$ is

$$A\mathbf{x} = \begin{bmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n \\ \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n.$$

This is precisely the coefficient set of a linear system, also written as a linear combination of the columns $\mathbf{a}_i$ of $A$. Hence if $\mathbf{b}$ is the column vector of constants, any linear system is equivalent to solving $A\mathbf{x} = \mathbf{b}$.

**Example 3.30**

Determine the product $A\mathbf{s}$ where

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}, \qquad \mathbf{s} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}.$$

*Solution.* Using matrix multiplication we get

$$A\mathbf{s} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1+1+2 \\ 1+2+6 \\ 2+3+2 \end{bmatrix} = \begin{bmatrix} 4 \\ 9 \\ 7 \end{bmatrix}.$$

Setting $\mathbf{b} = \begin{bmatrix} 4 & 9 & 7 \end{bmatrix}^T$, then this is precisely the statement that $\mathbf{s}$ is the solution to the linear system given in Example 3.9. ■

Matrix multiplication satisfies many familiar properties of multiplication. However, it also satisfies some very unfamiliar properties. For example, it is possible for $A$ to be a non-zero matrix, and $\mathbf{v}$ to be a non-zero vector, but still have $A\mathbf{v}$ be the zero vector, as evidenced by the following product :

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This is emblematic of a deeper problem: We can have $A\mathbf{v} = A\mathbf{w}$ but $\mathbf{v} \neq \mathbf{w}$. For example

$$\begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Additionally, matrix multiplication is not *commutative*; that is, generally $AB \neq BA$. To see this, let

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$$

for which

$$\begin{aligned} AB &= \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} \\ BA &= \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -3 \\ 1 & 0 \end{bmatrix}. \end{aligned}$$

These are not even close to being the same matrix. Finally, powers of non-zero matrices can be zero. For example, if

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

then $A$ is certainly not the 0 matrix, but

$$A^2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

so we have $A^2 = 0$.

## 3.5 Matrix Inversion

We use inversion to reverse an operation. For example, given the equation $ax = b$ for $a \neq 0$, to solve for $x$ we multiply both sides by $a^{-1}$ to get

$$a^{-1}ax = a^{-1}b \quad \Rightarrow \quad x = a^{-1}b.$$

We would like to do something similar for matrices.

**Definition 3.31**

Let $A$ be a $n \times n$ matrix. We say that $A$ is *invertible* with inverse $B$ if

$$AB = I_n, \qquad BA = I_n.$$

We often denote the inverse of $A$ by $A^{-1}$. This does precisely what we want in terms of solving linear systems: Given an linear system $A\mathbf{x} = \mathbf{b}$ such that $A$ is invertible, we can apply its inverse $A^{-1}$ to both sides to get

$$A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} \quad \Rightarrow \quad \mathbf{x} = A^{-1}\mathbf{b}.$$

However, unlike real numbers, not all non-zero matrices have inverses. For example, you can show that the matrix

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

does not have an inverse by explicitly trying to compute one.

In the special case of $2 \times 2$ matrices, the inverse is given by

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Rightarrow \quad A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}. \tag{3.3}$$

We can check my multiplying:

$$AA^{-1} = \frac{1}{ad-bc} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad-bc} \begin{bmatrix} ad-bc & -ab+ab \\ cd-cd & -bc+ad \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

with $A^{-1}A$ similar.

Notice we cannot apply (3.3) to the matrix

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

which I told you was not invertible. Indeed $ad - bc = 0$, meaning we'd have to divide by zero. It turns out that a $2 \times 2$ matrix is invertible if and only if $ad - bc \neq 0$. This generalizes to something known as the *determinant*, which we will discuss in Section 3.6.

**Example 3.32**

Solve the linear system $A\mathbf{x} = \mathbf{b}$ if

$$A = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} -7 \\ 5 \end{bmatrix}.$$

*Solution.* This is the same linear system as Example 3.8, and there we found the solution $(x_1, x_2) = (1, 3)$. By (3.3) the inverse of $A$ is given by

$$A^{-1} = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}.$$

Applying this to $A\mathbf{x} = \mathbf{b}$ to solve for $\mathbf{x}$, we get

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -7 \\ 5 \end{bmatrix} = \begin{bmatrix} -14+15 \\ -7+10 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}. \qquad \blacksquare$$

There are formulas for inverting $3 \times 3$ and higher matrices, but in general they are too messy to be worth remembering. Instead, let $\mathbf{e}_i$ be the standard basis for $\mathbb{R}^n$, and write the columns of $A^{-1}$ as $\mathbf{f}_i$. The equation $AA^{-1} = I_n$ is equivalent to

$$\begin{aligned} I_n &= \begin{bmatrix} \mathbf{e}_1 & \ldots & \mathbf{e}_n \end{bmatrix} \\ &= AA^{-1} \\ &= A \begin{bmatrix} \mathbf{f}_1 & \cdots & \mathbf{f}_n \end{bmatrix} \\ &= \begin{bmatrix} A\mathbf{f}_1 & \cdots & A\mathbf{f}_n \end{bmatrix}. \end{aligned}$$

By equating, we want to solve the linear system $A\mathbf{f}_i = \mathbf{e}_i$ to find the $\mathbf{f}_i$. We know we can do this with the augmented matrix $\left[A \mid \mathbf{f}_i\right]$, but rather than have to do this for every $\mathbf{f}_i$, we can do them all simultaneously by using the augmented matrix

$$\left[\begin{array}{c|cccc} A & \mathbf{f}_1 & \mathbf{f}_2 & \cdots & \mathbf{f}_n \end{array}\right].$$

If the left portion of the augmented matrix cannot be reduced to the identity matrix, then the matrix is not invertible.

**Example 3.33**

Find $A^{-1}$ and use it to solve the linear system $A\mathbf{x} = \mathbf{b}$ where

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 4 \\ 9 \\ 7 \end{bmatrix}.$$

*Solution.* This is the same linear system given in Example 3.9, where we found a solution of $(x_1, x_2, x_3) = (1, 1, 2)$. Setting up our augmented system and row reducing, we get

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 1 & 0 \\ 2 & 3 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[(-2)R_1+R_3\to R_3]{(-1)R_1+R_2\to R_2} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 1 & -1 & -2 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{(-1)R_2+R_3\to R_3} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 0 & -3 & -1 & -1 & 1 \end{array}\right] \xrightarrow{(-1/3)R_3\to R_3} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1/3 & 1/3 & -1/3 \end{array}\right]$$

$$\xrightarrow[(-1)R_3+R_1\to R_1]{(-2)R_3+R_2\to R_2} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 2/3 & 1/3 & 1/3 \\ 0 & 1 & 0 & -5/3 & 1/3 & 2/3 \\ 0 & 0 & 1 & 1/3 & 1/3 & -1/3 \end{array}\right] \xrightarrow{(-1)R_2+R_1\to R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 7/3 & -2/3 & -1/3 \\ 0 & 1 & 0 & -5/3 & 1/3 & 2/3 \\ 0 & 0 & 1 & 1/3 & 1/3 & -1/3 \end{array}\right]$$

For simplicity, we factor out the 1/3 term and write

$$A^{-1} = \frac{1}{3} \begin{bmatrix} 7 & -2 & -1 \\ -5 & 1 & 2 \\ 1 & 1 & -1 \end{bmatrix}.$$

We can then solve the linear system as

$$\mathbf{x} = A^{-1}\mathbf{b} = \frac{1}{3}\begin{bmatrix} 7 & -2 & -1 \\ -5 & 1 & 2 \\ 1 & 1 & -1 \end{bmatrix}\begin{bmatrix} 4 \\ 9 \\ 7 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 28-18-7 \\ -20+9+14 \\ 4+9-7 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix},$$

which agrees with what we found earlier. ■

Example 3.33 was significantly more difficult than Example 3.9 where we just used row reduction. Why then would we ever want to compute the inverse? The problem with row reduction is that, were we to change the constants in $\mathbf{b}$, we would have to do the entire row reduction over again. On the other hand, computing the inverse is a one-time thing. Once you have it, you can quickly solve $A\mathbf{x} = \mathbf{b}$ for any $\mathbf{b}$. So it depends on whether you need to solve $A\mathbf{x} = \mathbf{b}$ for many different $\mathbf{b}$.

**Theorem 3.34**

Suppose that each $A_i$ is an invertible $n \times n$ matrix.

1. $(A^{-1})^{-1} = A$
2. $(A_1A_2\cdots A_k)^{-1} = A_k^{-1}\cdots A_2^{-1}A_1^{-1}$
3. $(A^k)^{-1} = (A^{-1})^k$ for all $k$
4. $(cA)^{-1} = (1/c)A^{-1}$ for $c \neq 0$
5. $A$ invertible if and only if $A^T$ invertible.
6. $(A^{-1})^T = (A^T)^{-1}$.

**Remark 3.35**

1. Computing inverses using Gaussian elimination is actually a bad way of computing inverses. Modern computers use more sophisticated techniques to compute inverses.
2. Almost every $n \times n$ matrix is invertible. What I mean by this is that if you created an $n \times n$ matrix by randomly choosing the entries, it would be mathematically impossible for you to create a non-invertible matrix. The word 'random' here is important though. Certainly we can construct non-invertible matrices if we are allowed to choose the entries within the matrix.

**Example 3.36**

Suppose that

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}, \quad A^{-1} = \frac{1}{3}\begin{bmatrix} 7 & -2 & -1 \\ -5 & 1 & 2 \\ 1 & 1 & -1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

If $B = ADA^{-1}$, compute $B^{-1}$ and $B^2 + B$.

*Solution.* Using brute force, you could explicitly compute $B$, then apply our algorithm above for computing the inverse, but this is a lot of work. Using our properties of inversion, we can simplify the process. For example,

$$B^{-1} = (ADA^{-1})^{-1} = (A^{-1})^{-1}D^{-1}A^{-1} = AD^{-1}A^{-1}.$$

Since $D$ is a diagonal matrix, its inverse is just the reciprocal of the diagonal entries, so

$$\begin{aligned} B^{-1} = AD^{-1}A^{-1} &= \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}\begin{bmatrix} 7 & -2 & -1 \\ -5 & 1 & 2 \\ 1 & 1 & -1 \end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix} 1 & 1/2 & 1/3 \\ 1 & 1 & 1 \\ 2 & 3/2 & 1/3 \end{bmatrix}\begin{bmatrix} 7 & -2 & -1 \\ -5 & 1 & 2 \\ 1 & 1 & -1 \end{bmatrix} = \frac{1}{18}\begin{bmatrix} 6 & 3 & 2 \\ 6 & 6 & 6 \\ 12 & 9 & 2 \end{bmatrix}\begin{bmatrix} 7 & -2 & -1 \\ -5 & 1 & 2 \\ 1 & 1 & -1 \end{bmatrix} \\ &= \frac{1}{18}\begin{bmatrix} 29 & -7 & -2 \\ 18 & 0 & 0 \\ 41 & -13 & 4 \end{bmatrix}. \end{aligned}$$

Similarly, note that

$$B^2 = (ADA^{-1})^2 = (ADA^{-1})(ADA^{-1}) = AD^2A^{-1}$$

with $D^2$ computed easily as the square of the elements on the diagonal. Thus

$$B^2 + B = AD^2A^{-1} + ADA^{-1} = A(D^2 + D)A^{-1},$$

which can be computed as

$$B^2 + B = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 12 \end{bmatrix}\begin{bmatrix} 7 & -2 & -1 \\ -5 & 1 & 2 \\ 1 & 1 & -1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} -4 & 14 & -2 \\ -10 & 44 & -14 \\ -50 & 22 & 20 \end{bmatrix}. \qquad \blacksquare$$

## 3.6 Determinants

In this section we analyze the *determinant of a matrix*. Very loosely, the determinant is map which assigns to each matrix a real-number. The value of this real number has several interpretations. Sometimes we care about the magnitude of this number, sometimes the sign, and sometimes we are only interested in whether the number is non-zero. For example, the determinant will give us a way of determining whether a matrix is invertible, without having to explicitly compute the inverse.

Unfortunately, most of the ways of writing down the determinant are complicated. The definitions which are theoretically useful are poor for computation, and the definitions which are useful for computation are poor theoretically. Even those which are computationally valuable turn out to be resource intensive.

### 3.6.1 Definition

As mentioned above, the determinant map which assigns to each matrix a real number. The definition we will use for the determinant will be by *cofactor expansion*, alternatively known as the

*Laplace expansion.* To begin with, if $A = [a]$ is a $1 \times 1$ matrix then its determinant is $\det(A) = a$. If $A$ is a $2 \times 2$ matrix, its determinant is defined to be

$$\det(A) = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

the product of the diagonal minus the product of the anti-diagonal. The $3 \times 3$ case is trickier. I will write it down, then comment on precisely how it is calculated. Let

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

for which

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \tag{3.4}$$

This looks esoteric and arbitrary, but there is a method to the madness.

**Definition 3.37**

Let $A$ be an $n \times n$ matrix. For any $1 \leq i, j \leq n$, the $(i, j)$-*submatrix* of $A$, denoted $M_{ij}$, is the $(n-1) \times (n-1)$ matrix formed by deleting the $i$th row and $j$th column from $A$. The $(i, j)$-cofactor of $A$, denoted $C_{ij}$, is $C_{ij} = (-1)^{i+j} \det(M_{ij})$.

**Example 3.38**

Determine the $(1, 3)$- and $(2, 3)$-cofactor of $A = \begin{bmatrix} 1 & 4 & -2 \\ 3 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$.

*Solution.* The $(1, 3)$-cofactor is $C_{13} = (-1)^{1+3} \det(M_{13}) = \det(M_{13})$ where $M_{13}$ is the submatrix formed by deleting the first row and third column of $A$, hence

$$C_{13} = \det \begin{bmatrix} 3 & -1 \\ 0 & 1 \end{bmatrix} = (3 \times 1) - (-1 \times 0) = 3.$$

Similarly, the $(2, 3)$-cofactor is $C_{23} = (-1)^{2+3} \det(M_{23}) = -\det(M_{23})$, so

$$C_{23} = -\det \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix} = -\left[(1 \times 1) - (4 \times 0)\right] = -1. \qquad \blacksquare$$

Notice we can write (3.4) as

$$\begin{aligned} \det(A) &= a_{11}\left(a_{22}a_{33} - a_{23}a_{32}\right) - a_{12}\left(a_{21}a_{33} - a_{23}a_{32}\right) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\ &= a_{11} \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} + a_{21}(-1) \det \begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix} + a_{13} \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \\ &= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}. \end{aligned}$$

That is, the determinant of a $3 \times 3$ matrix is a weighted sum of the cofactors along the first row of the matrix! I do not expect you to have guessed this was the case, and instead this is what we'll use for the definition of the determinant. An important point, is that there is nothing special about the first row. We could use any other row or column. For example, you can check that

$$\det(A) = a_{12}C_{12} + a_{22}C_{22} + a_{23}C_{23}$$

yields exactly the same formula as (3.4), where now we have done a weighted sum of cofactors along the first column.

**Definition 3.39**

If $A$ is an $n \times n$ matrix, then the *determinant* of $A$ is the weighted sum of the cofactors along any row or column. For example, along the $i$th row or $j$th column:

$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}.$$

There is no reason you should believe this quantity is invariant of choice of row or column, but the proof is horrific using the definition given, and so is omitted.

**Example 3.40**

Compute $\det \begin{bmatrix} 1 & 4 & -2 \\ 3 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$.

*Solution.* Since we have already computed the cofactors $C_{13}$ and $C_{23}$, it makes most sense to perform our cofactor expansion along the third column. To do this we need to determine $C_{33}$, which computation yields

$$C_{33} = (-1)^{3+3}\det(M_{33}) = \det \begin{bmatrix} 1 & 4 \\ 3 & -1 \end{bmatrix} = -1 - 12 = -13.$$

Putting this all together, we get

$$\det(A) = a_{13}C_{13} + a_{23}C_{23} + a_{33}C_{33} = (-2 \cdot 3) + (0 \cdot -1) + (1 \times -13) = -19. \qquad \blacksquare$$

Notice how the presence of a zero in the $(2, 3)$-position made our lives easier? As a general rule, if computing the derivative via cofactor expansion, it makes the most sense to expand along the row/column which contains the most zeroes. In fact, if a matrix has a row or column consisting entirely of zeroes, cofactor expansion along that row/column will always yield a determinant of 0. Of note is that the identity matrix $I_n$ has determinant $\det(I_n) = 1$ for any $n$.

**Example 3.41**

Compute the determinant of $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 2 & 2 & 2 \end{bmatrix}$.

*Solution.* Expanding over the first column (since it has the most zeroes), we get

$$\begin{aligned}\det(A) &= (-1)^{1+1}(1)\det\begin{bmatrix}-1 & -2\\ 2 & 2\end{bmatrix} + (-1)^{1+2}(0)\det\begin{bmatrix}2 & 3\\ 2 & 2\end{bmatrix} + (-1)^{1+3}(2)\det\begin{bmatrix}2 & 3\\ -1 & -2\end{bmatrix}\\ &= [(-2)-(-4)] + 2\,[(-4)-(-3)]\\ &= 0.\end{aligned}$$

Let's see that this answer is the same if expanded across the second row instead. Here we would get

$$\begin{aligned}\det(A) &= (-1)^{2+1}(0)\det\begin{bmatrix}2 & 3\\ 2 & 2\end{bmatrix} + (-1)^{2+2}(-1)\det\begin{bmatrix}1 & 3\\ 2 & 2\end{bmatrix} + (-1)^{2+3}(-2)\det\begin{bmatrix}1 & 2\\ 2 & 2\end{bmatrix}\\ &= (-1)\,[(2)-(6)] + 2\,[(2)-(4)]\\ &= 0.\end{aligned}$$

■

**Exercise:** Compute the determinant of the matrix given in Example 3.41 by expanding along any other row or column, and check to make sure that you got the same answer as computed above.

### 3.6.2 Properties of the Determinant

Dealing with determinants can be a big pain, so we would like to develop some tools to make our lives a little bit easier. The most useful tool will be the following:

**Theorem 3.42**

If $A, B$ are two $n \times n$ matrices, then $\det(AB) = \det(A)\det(B)$.

Notice the curious fact that the determinant does not care about the order of multiplication, since the product on the right-hand side $\det(A)\det(B) = \det(B)\det(A)$ is an operation in $\mathbb{R}$. We omit the proof of this theorem, but let us compute a few examples to check its veracity.

**Example 3.43**

Let $A = \begin{bmatrix}1 & 5\\ 2 & 3\end{bmatrix}$ and $B = \begin{bmatrix}-1 & 1\\ 5 & -2\end{bmatrix}$. Determine $\det(A)$, $\det(B)$, $\det(AB)$, and $\det(BA)$.

*Solution.* Straightforward computation yields

$$\det(A) = -7, \quad \det(B) = -3.$$

The product matrices are

$$AB = \begin{bmatrix}1 & 5\\ 2 & 3\end{bmatrix}\begin{bmatrix}-1 & 1\\ 5 & -2\end{bmatrix} = \begin{bmatrix}24 & -9\\ 13 & -4\end{bmatrix}, \qquad \det(AB) = (-96+117) = 21 = \det(A)\det(B),$$

$$BA = \begin{bmatrix} -1 & 1 \\ 5 & -2 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 1 & 19 \end{bmatrix}, \qquad \det(BA) = 19 + 2 = 21 = \det(A)\det(B). \qquad \blacksquare$$

Note however that the determinant is **not** additive; that is, $\det(A+B) \neq \det(A) + \det(B)$. Indeed, almost any pair of matrices will break this. A simple example is to take $A = I_2$ and $B = -I_2$. Then $A+B$ is the zero matrix, so $\det(A+B) = 0$. On the other hand, $\det(A) + \det(B) = 1 + 1 = 2$.

**Exercise:** Show that $\det(AB) = \det(A)\det(B)$ explicitly in the $2 \times 2$ case.

**Corollary 3.44**

If $A$ is an invertible $n \times n$ matrix, then $\det(A^{-1}) = 1/\det(A)$.

*Proof.* We know that $AA^{-1} = I_n$, so applying the determinant we have $\det(AA^{-1}) = \det(A)\det(A^{-1}) = \det(I_n) = 1$. Isolating for $\det(A^{-1})$ we get

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

as required. $\square$

**Proposition 3.45**

If $A$ is an $n \times n$ matrix, then $\det(A) = \det(A^T)$

## 3.7 Eigenvalues and Eigenvectors

Eigenvalues and eigenvectors are one of the most important applications of linear algebra, since there is a sense in which a matrix is effectively determined by these values. The word 'eigen' comes from the German word 'own,' as in 'belong to.'

**Definition 3.46**

Let $A$ be an $n \times n$ matrix. A (real) *eigenvalue* of $A$ is a $\lambda \in \mathbb{R}$ such that there exists a non-zero vector $\mathbf{v}_\lambda$ satisfying

$$A\mathbf{v}_\lambda = \lambda \mathbf{v}_\lambda.$$

In such an instance, we say that $\mathbf{v}_\lambda$ is an *eigenvector* of $A$ corresponding to the eigenvalue $\lambda$.

For example,

$$\begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -2 \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

so $\lambda = -2$ is an eigenvalue of this matrix with associated eigenvector $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$. Notice if we substitute $\begin{bmatrix} 4 & 4 \end{bmatrix}^T$ we would get

$$\begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 8 - 16 \\ -4 - 4 \end{bmatrix} = -2 \begin{bmatrix} 4 \\ 4 \end{bmatrix},$$

go that $\begin{bmatrix} 4 & 4 \end{bmatrix}^T$ is also an eigenvector with the same eigenvalue $\lambda = -2$. Interesting! There's nothing special about the number 4 here. More generally, if $\mathbf{v}_\lambda$ is an eigenvector of $A$ with eigenvalue $\lambda$, and $c \in \mathbb{R}$, then

$$A(c\mathbf{v}_\lambda) = c(A\mathbf{v}_\lambda) = c(\lambda \mathbf{v}_\lambda) = \lambda(c\mathbf{v}_\lambda).$$

This shows that $c\mathbf{v}_\lambda$ is also eigenvector of $A$ with eigenvalue $\lambda$.

So how do we find eigenvalues and eigenvectors? Recognize that we can re-write $A\mathbf{v}_\lambda = \lambda\mathbf{v}_\lambda$ as $(A - \lambda I)\mathbf{v}_\lambda = 0$. In particular, we are asking that the matrix $(A - \lambda I)$ send a non-zero vector $\mathbf{v}_\lambda$ to the zero vector. This can only happen if $(A - \lambda I)$ is *not* invertible; that is, if $\det(A - \lambda I) = 0$. Computing $\det(A - \lambda I)$ will result in a polynomial in the variable $\lambda$, known as the *characteristic polynomial*. If we can find the roots of this polynomial, we will have the eigenvalues.

So let's compute the determinant of $A - \lambda I$ and see what we get. If $A = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}$ as above then

$$\begin{aligned} 0 = \det(A - \lambda I) &= \det \begin{bmatrix} 2 - \lambda & -4 \\ -1 & -1 - \lambda \end{bmatrix} \\ &= (2 - \lambda)(-1 - \lambda) - 4 = \lambda^2 - \lambda - 6 \\ &= (\lambda - 3)(\lambda + 2). \end{aligned}$$

which is zero when $\lambda = 3$ and $\lambda = -2$. We already knew that $\lambda = -2$ via the example above, but now we see that there is another eigenvalue at $\lambda = 3$. Let's compute the eigenvector associated to $\lambda = 3$. We know that $(A - 3I)\mathbf{v}_3 = 0$, so if $\mathbf{v}_3 = (v_1, v_2)^T$ we get

$$\begin{aligned} (A - 3I)\mathbf{v}_3 &= \begin{bmatrix} -1 & -4 \\ -1 & -4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\ &= \begin{bmatrix} -v_1 - 4v_2 \\ -v_1 - 4v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \end{aligned}$$

Both equations give the same information, so just looking at one of them we have $v_1 = -4v_2$. This means that any vector which looks like

$$\begin{bmatrix} -4v_2 \\ v_2 \end{bmatrix} = \begin{bmatrix} -4 \\ -1 \end{bmatrix} v_2$$

will be an eigenvector for $\lambda = 3$. A simple choice might be to set $v_2 = 1$, so that $\mathbf{v}_3 = \begin{bmatrix} -4 & 1 \end{bmatrix}^T$. You can check that

$$\begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} -4 \\ 1 \end{bmatrix} = 3 \begin{bmatrix} -4 \\ 1 \end{bmatrix}.$$

Note that sometimes eigenvalues might not exist, for example, if we try to compute that eigenvalues of the matrix $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ we get

$$\det(A - \lambda I) = \det \begin{bmatrix} -\lambda & 1 \\ -1 & -\lambda \end{bmatrix} = \lambda^2 + 1$$

which has no roots.

**Example 3.47**

Compute the eigenvalues and eigenvectors for the $3 \times 3$ matrix

$$A = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}.$$

*Solution.* We have

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{bmatrix} \\ &= (1-\lambda)\left[(2-\lambda)(1-\lambda) - 1\right] - (1-\lambda) \\ &= (1-\lambda)\left[(2-\lambda)(1-\lambda) - 2\right] \\ &= (1-\lambda)\left[\lambda^2 - 3\lambda + 2 - 2\right] \\ &= \lambda(1-\lambda)(\lambda - 3). \end{aligned}$$

Hence our eigenvalues are $0, 1, 3$. When $\lambda = 0$ we row reduce to find

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \xrightarrow{R_1 + R_2 \to R_2} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \xrightarrow{R_2 + R_3 \to R_3} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence if $\mathbf{v} = (v_1, v_2, v_3)$ then $v_1 = v_2 = v_3$. A nice choice is $(1, 1, 1)$. When $\lambda = 1$ we have

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \end{bmatrix} \xrightarrow[R_1 \leftrightarrow R_2]{-R_1 + R_3 \to R_3} \begin{bmatrix} -1 & 1 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence our eigenvector is $(1, 0, -1)$. Finally, if $\lambda = 3$ then

$$\begin{bmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & -2 \end{bmatrix} \xrightarrow[R_2 \leftrightarrow R_1]{-2R_2 + R_1 \to R_1} \begin{bmatrix} -1 & -1 & -1 \\ 0 & 1 & 2 \\ 0 & -1 & -2 \end{bmatrix} \xrightarrow{R_2 + R_2 \to R_2} \begin{bmatrix} -1 & -1 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

yielding an eigenvector $(1, -2, 1)$. ■

## 3.8 Exercises

3-1. Solve each given system of equations:

(a) $\begin{aligned} 7x + \quad 2 &= -19 \\ -x + 2y &= \quad 21 \end{aligned}$

(b) $\begin{aligned} -8x - 10 &= 24 \\ 6 + 5y &= \quad 2 \end{aligned}$

(c) $\begin{aligned} 2x + \quad y + 3z &= \quad 1 \\ 4x + 5y + 7z &= \quad 7 \\ 2x - 5y + 5z &= -7 \end{aligned}$

(d) $\begin{aligned} x + 2y + 4z &= 1 \\ x + \quad y + 3z &= 2 \\ 2x + 5y + 9z &= 1 \end{aligned}$

(e) $\begin{aligned} x - 3z &= \quad 8 \\ 2x + 2y + \ 9z &= 7 \\ y + 5z &= -2 \end{aligned}$

(f) $\begin{aligned} x - 3y &= 5 \\ -x + \quad y \ + \ 5z &= 2 \\ y + \quad z &= \quad 0 \end{aligned}$

(g) $\begin{aligned} y + 4z &= -5 \\ x + 3y + \ 5z &= -2 \\ 3x + 7y + \ 7z &= \quad 6 \end{aligned}$

(h) $\begin{aligned} 2x - 3y + \quad z + 7w &= \quad 14 \\ 2x + 8y - 4z + 5w &= \quad -1 \\ x + 3y - 3z &= \quad 4 \\ -5x + 2y + 3z + 4w &= -19 \end{aligned}$

(i) $\begin{aligned} 2x + 4y + 5z + \quad 7w &= -26 \\ x + 2y + \quad z - \quad w &= \quad -4 \\ -2x - 4y + \quad z + 11w &= -10 \end{aligned}$

3-2. Determine if the following systems have no solutions, a unique solution, or infinitely many solutions. You do not need to solve the system.

(a) $\begin{aligned} 2x - 5y + 8z \quad &= 0 \\ -2x - 7y + z \quad &= 0 \\ 4x + 2y + 7z \quad &= 0 \end{aligned}$

(b) $\begin{aligned} x - 3y + 7z \quad &= 0 \\ -2x + y - 4z \quad &= 0 \\ x + 2y + 9z \quad &= 0 \end{aligned}$

(c) $\begin{aligned} 3x - 2y \quad &= 3 \\ 6x - 4y \quad &= 4 \end{aligned}$

3-3. Row reduce the following matrices

(a) $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 2 \\ 1 & 0 & 1 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 2 & 1 & 2 & 1 \\ 2 & 1 & 2 & 1 & 2 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 2 & 0 & 5 \\ 2 & 3 & 1 & 4 \\ -1 & -1 & -1 & 1 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 3 \\ 2 & -1 \\ -1 & -3 \end{pmatrix}$

3-4. Consider the linear system

$$\begin{aligned} x + \quad y - \qquad\quad z &= 10 \\ -x + \quad y - \qquad 11z &= \ 2 \\ 2x + 4y + (8+a)z &= 50 \end{aligned}$$

(a) Find the values of $a$ such that this system has no solutions.

(b) Find the values of $a$ such that this system has a unique solution. What is that solution?

(c) Find the values of $a$ such that this system has infinitely many solutions. What are those solutions?

3-5. Consider a closed economic system consisting of three industries: Telecommunications (T), Hydro Electric (H), and Service (S). Suppose these industries have the following exchange table:

| | Distribution of Output | | |
|---|---|---|---|
| Purchased By | (T) | (H) | (S) |
| (T) | 0.1 | 0.6 | 0.4 |
| (H) | 0.2 | 0 | 0.4 |
| (S) | 0.7 | 0.4 | 0.2 |

For example, column 1 describes the total economic output of the telecommunications industry. Of that output, 20% is purchased by the hydro electric industry, 70% is purchased by the service industry, and 10% is retained as operational costs.

Leontief's Economic Equilibrium states that there exist equilibrium prices that can be assigned to the total outputs of the various sectors in such a way that the income of each sector exactly balances the expenses.

(a) Let $p_T, p_H, p_S$ be the prices of the total annual output of the telecommunications, hydro electric, and service industries. Set up a system of equations which describes Leontief's equilibrium.

(b) Solve the system from part (a) to find the equilibrium prices. You may use a calculator or a computer algebra system for this. Round your answers to two decimal places.

3-6. You are preparing breakfast for an eccentric grandfather, who has lost his sense of taste, but insists that the nutritional value of his breakfast must consist of exactly

Calories: 625 kcal  Protein: 12 g  Carbs: 161 g  Fat: 5g.

Available in the house are four cereals: Cheerios, Frosted Flakes, Raisin Bran, and Shreddies. The following table gives the nutritional information per 25 gram serving

| | Nutritional Information per 25 gram serving | | | |
|---|---|---|---|---|
| Nutrient | Cheerios | Frosted Flakes | Raisin Bran | Shreddies |
| Calories (kcal) | 90 | 90 | 80 | 85 |
| Protein (g) | 3 | 1 | 2 | 2 |
| Carbs (g) | 26 | 22 | 28 | 21 |
| Fat (g) | 2 | 0 | 1 | 1 |

Determine a suitable mix of these cereals to accomplish a breakfast that matches his nutritional demands. How much cereal do you feed him?

3-7. You are a bouquet designer for the Grassmann Flower Company. You've come up with a dazzling new arrangement, which you are calling the "What in carnation!" bouquet. To fulfill an order for your new bouquet, you need

2100 orchids, 900 lilies, 2400 carnations, and 1200 hydrangeas.

Your company has four suppliers, each of which ships flowers in bunches:

| | |
|---|---|
| Alexey's Awesome Amaryllis: | 1 bunch = 2 orchids and 1 lilies. |
| David's Dazzling Daisies: | 1 bunch = 1 orchid and 1 carnation. |
| Srishti's Superb Snapdragons: | 1 bunch = 3 carnations and 1 hydrangeas. |
| Paula's Perfect Peonies: | 1 bunch = 3 orchids, 2 lilies, 2 hydrangeas. |

Determine, by any means available to you, how many bunches you should buy from each supplier in order to make the "What in carnation!" bouquet.

3-8. Define the matrix $A$ and vectors $\mathbf{b}_i, i = 1, 2, 3, 4$ as follows:

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 0 & 5 \\ -1 & 1 & 7 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} 2 \\ 9 \\ 7 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} 2 \\ 38 \\ 44 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} 52 \\ -55 \\ -67 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} 18 \\ -16 \\ -26 \end{bmatrix}.$$

Note that the steps used in reducing the augmented matrix $\left[\begin{array}{c|c} A & \mathbf{b}_i \end{array}\right]$ to RREF work regardless of which $\mathbf{b}_i$ is in the augmented column. Therefore, we could solve our linear system for all four choices of $\mathbf{b}_i$ by creating the super augmented matrix $\left[\begin{array}{c|cccc} A & \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 & \mathbf{b}_4 \end{array}\right]$ and putting this matrix into RREF. Try it! What is the RREF of the super augmented matrix?

3-9. Find the rank of each given matrix:

(a) $\begin{pmatrix} 1 & 2 & 3 & 4 \\ -3 & -2 & 1 & 1 \\ 8 & 8 & 4 & 6 \end{pmatrix}$

(b) $\begin{pmatrix} 2 & 0 & 1 & -1 \\ 0 & 1 & 2 & 1 \\ 2 & -1 & -1 & -2 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 2 & 4 & 2 \\ 0 & 2 & 2 & 1 \end{pmatrix}$

(d) $\begin{pmatrix} 2 & 2 & -1 & 1 \\ 4 & 0 & 2 & 2 \\ 0 & 6 & -3 & 4 \end{pmatrix}$

(e) $\begin{pmatrix} 0 & 16 & 8 & 4 \\ 2 & 4 & 8 & 16 \\ 16 & 8 & 4 & 2 \\ 4 & 8 & 16 & 2 \end{pmatrix}$

3-10. Let $A = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 2 & 3 & -6 \\ 1 & 0 & 0 \\ 3 & 3 & -1 \end{bmatrix}$. Determine the given expression.

(a) $A + B + C$

(b) $2A - 3B$

(c) $A + 3(B - C) - 2B$

(d) $A^T + B - 3C^T$

3-11. Let $A = \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix}, B = \begin{bmatrix} y & -y \\ 1 & 1 \end{bmatrix}$. Find the values of $x, y$ such that $2A - B = \begin{bmatrix} 7 & -2 \\ -1 & 4 \end{bmatrix}$

3-12. Determine whether $\mathbf{v}$ is a linear combination of $\mathbf{x}, \mathbf{y}, \mathbf{z}$.

(a) $\mathbf{v} = \begin{bmatrix} -7 \\ 5 \\ 2 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 3 \\ -3 \\ 6 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} -7 \\ 5 \\ -4 \end{bmatrix}, \mathbf{z} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$.

(b) $\mathbf{v} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}, \mathbf{z} = \begin{bmatrix} -5 \\ 4 \\ -1 \end{bmatrix}$.

(c) $\mathbf{x} = \mathbf{y} + \mathbf{z}, \mathbf{y} = \mathbf{z}, \mathbf{z} = 3\mathbf{v}$

3-13. Define the following vectors in $\mathbb{R}^3$:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

(a) If $\mathbf{v} \in \mathbb{R}^3$ is an arbitrary vector, write down the linear system which addresses whether $\mathbf{v}$ is a linear combination of $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$.

(b) Show that for any vector $\mathbf{v} \in \mathbb{R}^3$, we can write $\mathbf{v}$ as a linear combination of $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$.

3-14. Two vectors $\mathbf{u}, \mathbf{v}$ are said to be orthogonal if $\mathbf{u} \cdot \mathbf{v} = 0$. For each pair of vectors below, determine the orthogonal pairs.

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \mathbf{v}_3 = \frac{1}{\sqrt{11}} \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} \pi \\ 0 \\ -e \end{bmatrix}, \quad \mathbf{v}_5 = \begin{bmatrix} 32 \\ -10 \\ 32 \end{bmatrix}, \quad \mathbf{v}_6 = \begin{bmatrix} 6 \\ 6 \\ -4 \end{bmatrix}$$

3-15. Define the matrices $A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 2 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} -1 & -1 \\ 0 & 3 \\ -5 & 2 \end{bmatrix}$. Determine whether the following products make sense, and if so compute the product.

(a) $BA$

(b) $CB$

(c) $AC$

(d) $ABC$

(e) $CBA$

(f) $BAC - ACB$

(g) $AC^TB$

(h) $A^2 + B^2$

3-16. If $A = \begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix}$, find a non-zero matrix $B$ such $AB = \mathbf{0}$, where $\mathbf{0}$ is a zero matrix of whatever size you like.

3-17. Define the matrices

$$A = \begin{bmatrix} 1 & 5 & -4 & -3 \\ 3 & -12 & -6 & 0 \\ 0 & 14 & 7 & 12 \\ 4 & 8 & -3 & 10 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 7 & 17 \\ 3 & -3 & 0 \\ 1 & 14 & 6 \\ 4 & 4 & 2 \end{bmatrix}.$$

Find the $(4,2)$ element of the matrix $AB$ without computing the entire product.

3-18. Show that the matrix

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 3 \\ 2 & 1 & 3 \end{bmatrix}$$

is a root of the polynomial $p(x) = x^3 - 5x^2 - 5$; that is, $p(A) = A^3 - 5A^2 - 5 = 0$.

3-19. True or False:

(a) If $\mathbf{u}$ and $\mathbf{v}$ are column vectors in $\mathbb{R}^n$, then $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}$.

(b) If $A$ is a square matrix satisfying $A^2 = 0$, then $A = 0$.

(c) If $A$ is a square matrix satisfying $A^2 = A$, then $A = \pm I$ or $A = 0$.

(d) There is a square matrix $A$ (of any dimension) such that $A^2 = -I$.

(e) If $A$ and $B$ are invertible $n \times n$ matrices, then $AB^T$ is invertible.

(f) If $A$ and $B$ are non-invertible, then $A + B$ is non-invertible.

(g) The equation $A\mathbf{x} = \mathbf{0}$ has a solution for any matrix $A$.

3-20. Let $D$ and $P$ be the matrices

$$P = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 2 \\ 1 & 3 & -2 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

(a) Find $P^{-1}$.

(b) If $A = 7PDP^{-1}$, find $A^{-1}$. *Hint:* Use the definition of $A$ rather than computing $A$ and then its inverse.

3-21. We know that there are three row operations: transposing rows, scaling a row, and adding a multiple of one row to another. Find, if possible, matrices which simulate these actions. For example, if $A$ is an $m \times n$ matrix, find an $m \times m$ matrix $T_{ij}$ such that $T_{ij}A$ is the same matrix as $A$ but with rows $i$ and $j$ interchanged.

3-22. Define the matrix

$$C = \begin{bmatrix} c & 0 & 1 \\ 2c & c & 1 \\ c & c^2 & 0 \end{bmatrix}.$$

(a) Determine the values of $c$ such that $C$ is invertible.

(b) For those values of $c$ you found in part (a), find $C^{-1}$.

3-23. Find the inverse of each matrix, if it exists:

(a) $\begin{bmatrix} 8 & 5 \\ -7 & -5 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & -2 & -1 \\ -1 & 5 & 6 \\ 5 & -4 & 5 \end{bmatrix}$

(d) $\begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$

3-24. Consider the matrix $A = \begin{bmatrix} 1 & 3 & 0 \\ 0 & 2 & 0 \\ 4 & 0 & 5 \end{bmatrix}$.

(a) Compute $\det(A)$.

(b) Compute $A^{-1}$

(c) Find the solution to $A\mathbf{x} = \mathbf{b}$ if $\mathbf{b} = \begin{bmatrix} 5 \\ -10 \\ 10 \end{bmatrix}$.

3-25. Let $A$ and $\mathbf{b}_i, i = 1, \ldots, 4$ be as defined in Problem 3-8. Find $A^{-1}$ and use it to solve the equation $A\mathbf{x}_i = \mathbf{b}_i$ for each $i = 1, \ldots, 4$.

3-26. (a) Find a $2 \times 2$ matrix $A$ such that $A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}$ for all $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$.

(b) Find a $3 \times 3$ matrix $B$ such that $B \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y+z \\ x+z \\ x+y \end{bmatrix}$ for all $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$.

(c) Find a $3 \times 3$ matrix $C$ such that $C \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y \\ z \\ 0 \end{bmatrix}$ for all $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$.

3-27. Suppose $\det(A) = 7$ and $\det(B) = -2$. What is $\det(A^{-2}B^T)$?

3-28. (a) Let $I_n$ be the $n \times n$ identity matrix. What is $\det(-2I_n)$?

(b) If $A$ is an $n \times n$ matrix and $c \in \mathbb{R}$ is a realy number, formulate a hypothesis for $\det(cA)$ in terms of $c$ and $\det(A)$.

3-29. Suppose $a, b, c$ are real numbers. Find the determinant of the following matrices:

(a) $A = \begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix}$.

(b) $B = \begin{bmatrix} 1 & -1 & 2 \\ 1+a & -1+a & 2-a \\ 2b & -2b & 4b \end{bmatrix}$.

(c) $C = \begin{bmatrix} a & 0 & a \\ 0 & b & 0 \\ a & b & c \end{bmatrix}$.

3-30. Let $A = \begin{bmatrix} 7 & 5 \\ -10 & -8 \end{bmatrix}$.

(a) Find the eigenvalues and eigenvectors of $A$.

(b) Find the eigenvalues and eigenvectors of $A^2$.

3-31. In each case indicate whether the statement is true or false. Be sure to give a justification as to why you have chosen your answer.

(a) If $A, B$ are $n \times n$ matrices, then $\det(AB) = \det(BA)$.

(b) If $A, B$ are $n \times n$ matrices, then $\det(A + B) = \det(A) + \det(B)$.

(c) If $A$ is an $n \times n$ matrix and $A = A^2$ then $\det(A) = 1$.

3-32. Suppose that $A$ is a $2 \times 2$ matrix. You are told that

- $\mathbf{v} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ is an eigenvector of $A$ corresponding to the eigenvalue $\lambda = -1$,

- $\mathbf{w} = \begin{bmatrix} 5 \\ -7 \end{bmatrix}$ is an eigenvector of $A$ corresponding to the eigenvalue $\lambda = 2$.

(a) Write the vector $\mathbf{u} = \begin{bmatrix} 7 \\ -5 \end{bmatrix}$ as a linear combination of $\mathbf{v}$ and $\mathbf{w}$.

(b) Compute $A\mathbf{u}$.

3-33. (Eigenvalues/Eigenvectors) Recall that a matrix $A$ is invertible if and only if $\det(A) \neq 0$.

(a) Show that if $\lambda = 0$ of $A$, then its charactierstic polynomial satisfies $c_A(0) = \det(A) = 0$.

(b) Conclude that if $\lambda = 0$ is an eigenvalue of $A$, then $A$ is not invertible.

(c) Show that if $\mathbf{x} \in \mathbb{R}^n$ is a non-zero vector such that $A\mathbf{x} = \mathbf{0}$, then $A$ is non-invertible.

3-34. Find the eigenvalues and eigenvectors of the following matrices:

(a) $\begin{bmatrix} 2 & 7 \\ -1 & -6 \end{bmatrix}$

(b) $\begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & -1 \\ 5 & -3 \end{bmatrix}$

(d) $\begin{bmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{bmatrix}$

(e) $\begin{bmatrix} 1 & 3 & -3 \\ -3 & 7 & -3 \\ -6 & 6 & -2 \end{bmatrix}$

3-35. (a) True or False: If $\mathbf{v}_\lambda$ is an eigenvector of $A$ with eigenvalue $\lambda$, then $\mathbf{v}_\lambda$ is also an eigenvector of $A^2$

(b) True or False: If $\mathbf{v}_\lambda$ is an eigenvector of $A$ with eigenvalue $\lambda$ and $A$ is invertible, then $\mathbf{v}_\lambda$ is also an eigenvector of $A^{-1}$.

(c) It is known that the product of the eigenvalues of a square matrix is the determinant of that matrix. True or False: A matrix with a zero eigenvalue is always invertible.

(d) True or False: If $A\mathbf{v}_\lambda = \lambda\mathbf{v}_\lambda$, and $B$ is another $n \times n$ matrix satisfying $B\mathbf{v}_\lambda = \mu\mathbf{v}_\lambda$, then $\mathbf{v}_\lambda$ is an eigenvector for $A + B$.

# 4 Probability and Counting

Almost everything in life operates on a principle of probability and likelihoods, making a mathematical understanding of these concepts invaluable. Markets are inherently stochastic, and any further study in fields like quantitative finance, game theory, or economics, requires a level of comfort in these ideas. This section will cover the basics of discrete probability, with a view towards Markov chains towards the end. Markov chains are a method by which we can model simple stochastic processes, and use mathematics to evaluate their asymptotic nature.

## 4.1 Counting

The title of this section might sound patronizing, but we're going to learn how to count. The arguments made below are combinatorial in nature, and require you to really meditate on what they mean.

**Basic Counting:** Our first problem is to count the different ways distinct objects can come together to form tuples. Your friend has an ice-cream buffet at her birthday party, from which a dessert is formed by choosing one option from each of the following three lists:

1. **Container:** Sugar Cone (SC), Waffle Cone (WC), Bowl (Bo)
2. **Ice Cream:** Vanilla (Va), Chocolate (Ch)
3. **Topping:** Sprinkles (Sp), Fudge (Fu), Caramel (Ca).

Taste-buds aside, how many different desserts can you make? If we list the possibilities as a tuple, such as $(\text{SC}, \text{Va}, \text{Sp})$ for a sugar cone-vanilla-sprinkle dessert, we can enumerate all the possibilities:

1. $(\text{SC}, \text{Va}, \text{Sp})$
2. $(\text{SC}, \text{Va}, \text{Fu})$
3. $(\text{SC}, \text{Va}, \text{Ca})$
4. $(\text{SC}, \text{Ch}, \text{Sp})$
5. $(\text{SC}, \text{Ch}, \text{Fu})$
6. $(\text{SC}, \text{Ch}, \text{Ca})$
7. $(\text{WC}, \text{Va}, \text{Sp})$
8. $(\text{WC}, \text{Va}, \text{Fu})$
9. $(\text{WC}, \text{Va}, \text{Ca})$
10. $(\text{WC}, \text{Ch}, \text{Sp})$
11. $(\text{WC}, \text{Ch}, \text{Fu})$
12. $(\text{WC}, \text{Ch}, \text{Ca})$
13. $(\text{Bo}, \text{Va}, \text{Sp})$
14. $(\text{Bo}, \text{Va}, \text{Fu})$
15. $(\text{Bo}, \text{Va}, \text{Ca})$
16. $(\text{Bo}, \text{Ch}, \text{Sp})$
17. $(\text{Bo}, \text{Ch}, \text{Fu})$
18. $(\text{Bo}, \text{Ch}, \text{Ca})$

So there are 18 possible desserts. Alternatively, we could model this as the decision tree shown in Figure 4.1. We'd like a way of counting the total possibilities without having to go through the onerous process of writing them all out. In the list above, notice the first entry in each column is the same, either SC, WC, or Bo. If we were to add a fourth container we would have a fourth column, or if we removed an option, we'd reduce to two columns. This suggests that the total number of ways of choosing the dessert is

$$(\#\text{ Desserts}) = (\#\text{ Containers}) \times (\#\text{ of ways of choosing flavour and topping}).$$

But we can apply exactly the same reasoning to the flavour of ice-cream, showing that

$$(\text{Ways of choosing a flavour and topping}) = (\#\text{ Flavours}) \times (\#\text{ of ways of choosing a topping}).$$

And finally, the number of ways of choose a topping is the same as the number of toppings. Hence

$$(\#\text{ Desserts}) = (\#\text{ Containers}) \times (\#\text{ Flavours}) \times (\#\text{ Toppings}) = 3 \times 2 \times 3 = 18.$$

A similar argument can be made using the decision tree. One argues that each node has the same number of branches emanating from it, and hence the total number of terminal nodes is multiplicative in each of the prior nodes.

This argument can be applied to any collection of decisions, leading to our first counting principle:

**Basic Counting Principle:** If $S_1, \ldots, S_n$ are a collection of finite sets, the number of ways of choosing one element from each set is $|S_1| \times |S_2| \times \cdots \times |S_n|$.

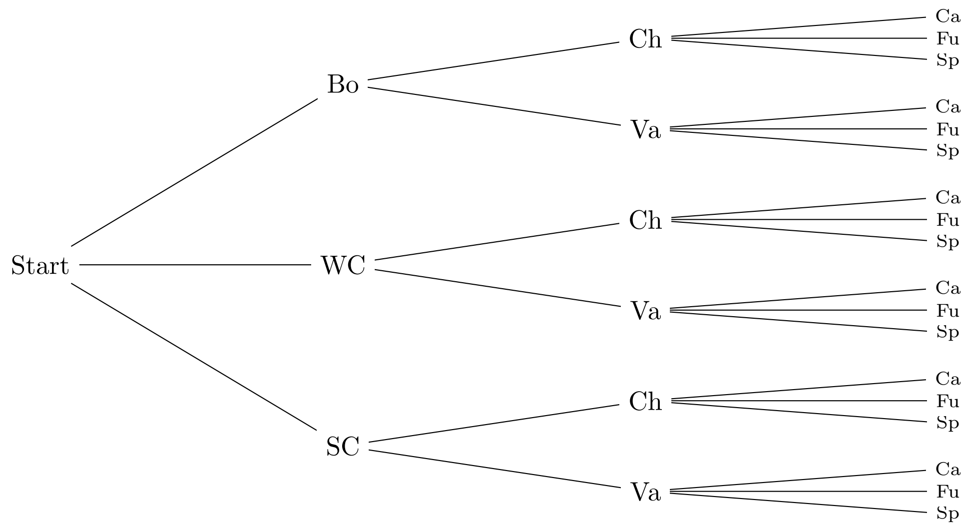



Figure 4.1: The decision tree for creating a dessert. Note that the number of branches attached to each node is independent of which node we choose.

**Example 4.1**

Suppose you flip a coin, throw a six-sided dice, and choose a letter of the alphabet. How many possible results are there?

*Solution.* A coin has two sides $S_c = \{H, T\}$, the dice has 6 faces $S_d = \{1, 2, 3, 4, 5, 6\}$, and the alphabet has 26 letters $S_a = \{a, b, c, \ldots, x, y, z\}$. The number of elements in each is

$$|S_c| = 2, \quad |S_d| = 6, \quad |S_a| = 26,$$

hence the total number of possible results is $2 \times 6 \times 26 = 312$. ■

**Example 4.2**

How many different ways are there to answer a multiple choice exam, consisting of 8 questions, each with 4 choices?

*Solution.* There are four ways of answering the first question, four ways of answering the second, and so on. Since there are eight total questions, the number of ways of answering the multiple choice exam is

$$\underbrace{4 \times 4 \times \cdots \times 4}_{\text{8-times}} = 4^8 = 65536. \qquad ■$$

**Example 4.3**

Suppose a set $S$ has $n$ elements. Determine the number of subsets of $S$.

*Solution.* This is a tricky but important example. Let's try a set and see what we get. For example, if $S = \{1, 2, 3\}$ then the subset of $S$ are

$$\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\},$$

of which there are $8 = 2^3$. We might guess that the number of subsets is therefore $2^n$. To see this using the Basic Counting Principal, think of the problem in the following manner: Line up the $n$ items and move through the items one at a time. In each case, you have the choice of either including this item in your subset, or not. Therefore each item has two possible configurations, and there are $n$-items, so the total number of subsets is

$$\underbrace{2 \times 2 \times \cdots \times 2}_{n\text{-times}} = 2^n. \qquad \blacksquare$$

**Permutations:** Next up are permutations. Given a finite set $S$ with $|S| = n$, a *$k$-permutation of $S$* is an ordered collection of $k$ distinct elements of $S$. For emphasis, *ordered* means that the order in which the elements are chosen matters, and *distinct* means that the elements must all be different.

For example, suppose the position of president, vice-president, and treasurer of the student union are to be randomly assigned to 3 different students. The candidate pool consists of 8 students, which we'll call $A, B, C, ..., H$. This is the problem of determining a 3-permutation from a set of size 8.

How many possible configurations are possible? Let's write a possible configuration as a concatenated triple, so that $ABC$ means $A$ is president, $B$ is vice-president, and $C$ is treasurer. A student cannot serve multiple positions, meaning something like $AAB$ is out of the question. Since order matters, $ABC$ is not the same as $CAB$. We can use the Basic Counting Principle, but we need to adapt our paradigm to this new situation. Suppose the president is chosen first, of which there are 8 possibilities. Once the president is chosen we move the vice-president, of which there are now 7 possibilities since one student has been removed from the pool. After this, the treasurer is chosen from the remaining 6 candidates. Hence there are

$$8 \times 7 \times 6 = 336$$

possible configurations of the student union.

The same argument works regardless of how many candidates there are originally, or how many positions need to be filled. This leads us to the following:

**Number of Permutations:** Let $S$ be a set consisting of $n$ elements, and $1 \leq k \leq n$ a natural number. The number of $k$-permutations of $S$ is

$${}_nP_k = (n)(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}.$$

If $k = 0$, we define ${}_nP_0 = 1$.

If you don't see the final equality, recall that $n! = (n)(n-1)(n-2)\cdots(3)(2)(1)$, so that

$$\begin{aligned}\frac{n!}{(n-k)!} &= \frac{(n)(n-1)(n-2)\cdots(n-k+2)(n-k+1)(n-k)!}{(n-k)!} \\ &= (n)(n-1)(n-2)\cdots(n-k+2)(n-k+1).\end{aligned}$$

**Example 4.4**

You are giving out a questionnaire asking students to rank their three favourite classes from

MAT133, ECO100, HIS101, RLG101, ANT101, and CIN101.

How many different choices are there?

*Solution.* Order matters, since the classes are being ranked. A student must choose three courses from the six listed, so there are

$$_6P_3 = \frac{6!}{3!} = 6 \times 5 \times 4 = 120$$

possible answers to the survey. ■

**Example 4.5**

How many distinct shuffles are there in a standard deck of 52 cards?

*Solution.* The order of the cards matter in a shuffle, but we're being asked to shuffle them all, so the number of distinct shuffles is $_{52}P_{52} = 52!$. This is a staggeringly big number: It is about $8 \times 10^{67}$. For context, the universe has been around $4.1 \times 10^{17}$ seconds, and there are about $2.4 \times 10^{67}$ atoms in the Milky Way Galaxy. ■

As mentioned, the number of permutations relies on the elements being distinct. However, there are occasions when we might want repeated objects. For example, the number of ways of arranging the letters "ARE" is 3!, but the number of ways of arranging "AREA" is not 4!, since we cannot distinguish between the two A's. Indeed, let's label the A's as $A_1$ and $A_2$ so we can see what happens.

$$\begin{array}{llllll}
A_1A_2ER, & A_1A_2RE, & A_1EA_2R, & A_1ERA_2, & A_1REA_2, & A_1RA_2\,E \\
A_2A_1ER, & A_2A_1RE, & A_2EA_1R, & A_2ERA_1, & A_2REA_1, & A_2RA_1\,E \\
EA_2A_1R, & EA_2RA_1, & EA_1A_2R, & EA_1RA_2, & ERA_1A_2, & ERA_2\,A_1 \\
RA_2EA_1, & RA_2A_1E, & REA_2A_1, & REA_1A_2, & RA_1EA_2, & RA_1A_2\,E
\end{array}$$

When we remove the indices, every word is counted twice. Hence the total number of arrangements of "AREA" is $4!/2 = 12$. What changes if we had used three A's? The answer is that every word would have been counted six times. For example, labelling $A_1$, $A_2$, and $A_3$, the word AREAA has the following 6 representations:

$$\mathrm{A_1REA_2A_3},\quad \mathrm{A_1REA_3A_2},\quad \mathrm{A_2REA_1A_3},\quad \mathrm{A_2REA_3A_1},\quad \mathrm{A_3REA_1A_2},\quad \mathrm{A_3REA_2A_1}.$$

That is, there are as many extra words as there are ways of arranging the $\mathrm{A}_1$, $\mathrm{A}_2$, and $A_3$; namely, 3!. The number of permutations is thus the total possible with labelling (5!), dividing by the number of times each number is over counted (3!).

If more than one letter is repeated, the same argument can be made, removing the duplicate counts by dividing by the number of possible configurations of the second letter, then the third, and so on.

**Number of Permutations with Repetition:** Suppose a collection of $n$ objects is given, of which $k$ are distinct. Suppose there are $n_i$ objects of the $i$th type, for $i \in \{1, \ldots, k\}$. The number of possible permutations of these $n$ objects is

$$\frac{n!}{n_1!n_2!n_3!\cdots n_{k-1}!n_k!}.$$

Changing paradigm slightly, the same formula can be used to determine the number of ways of classifying objects. For example, suppose you're organizing a field trip and have 10 children you need to take to the zoo. You have three vehicles, a car (C) which seats three, a mini-van (V) which seats 5, and a truck (T) which seats 2. The number of ways of sorting children into vehicles is

$$\frac{10!}{3!5!2!}.$$

To see this, consider an arrangement of the letters CCCVVVVVTT, which means that the first three children take the car, the next five take the van, and the last two take the truck. The number of ways of sorting children into cars is then equivalent to the number of permutations of these letters.

**Combinations:** Combinations correspond to those permutations where order does not matter. For example, the draw for the Lotto 6/49 consists of 6 balls labelled from 1 to 49. If your lottery ticket matches these numbers, you win. The order of the balls does not matter, so long as you have the same 6 numbers.

To determine the number of combinations, we'll count the number of permutations, and divide out the number of ways of rearranging each group with the same choices. The Lotto 6/49 is a bit much, so let's use the Lotto 3/4 for our example, wherein three balls are chosen from those numbered one to four. If order matters, we know there are ${}_4P_3 = 24$ possible choices:

| | | | |
|---|---|---|---|
| 123 | 124 | 134 | 234 |
| 132 | 142 | 143 | 243 |
| 213 | 214 | 314 | 324 |
| 231 | 241 | 341 | 342 |
| 312 | 412 | 413 | 423 |
| 321 | 421 | 431 | 432 |

Each column consists of the same three numbers, albeit in a different order. We want to count how many different ways there are of arranging the numbers. But we already know this! The number of ways of arranging three numbers is 3!, so we divide the total number of permutation ${}_4P_3$ by the number of ways of arranging three numbers 3! to get $4!/(1!3!) = 4$ different combinations.

**Number of Combinations:** Suppose a collection of $n$ objects is given, and $1 \leq k \leq n$ is a natural number. The number of $k$-combinations of those $n$ objects is

$${}_nC_k = \frac{{}_nP_k}{k!} = \frac{n!}{k!(n-k)!}.$$

When $k = 0$, we take ${}_nC_0 = 1$.

**Example 4.6**

Determine the number of Lotto 6/49 winning combinations.

*Solution.* A winning combination is any combination of the 6 numbers from 1 to 49, thus there are

$${}_{49}C_6 = \frac{49!}{6!43!} = \frac{49 \cdot 48 \cdot 47 \cdot 46 \cdot 45 \cdot 44}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 13,983,816$$

winning combinations. ■

From Example 4.3, we know that the total number of subsets of a set of size $n$ is $2^n$. Note that ${}_nC_k$ describes the total number of sets of size $k$, and therefore

$${}_nC_0 + {}_nC_1 + {}_nC_2 + \cdots {}_nC_{n-1} + {}_nC_n = \sum_{k=0}^{n} {}_nC_k = 2^n.$$

## 4.2 First Principles of Probability

Now that we know how to count, we can start looking at probabilities.

**Definition 4.7**

Given an experiment, the *sample space* of that experiment $S = \{s_1, \ldots, s_n\}$ is the collection of all possible outcomes. A *probability distribution on $S$* is a function $p$ on $S$ such that $p(s_i) \in [0,1]$ describes the probability that $s_i$ is the outcome of the experiment, and

$$\sum_{i=1}^{n} p(s_i) = 1.$$

If $p(s_1) = p(s_2) = \cdots = p(s_n) = 1/n$, then the experiment is said to be *equiprobable*. An *event* is a subset of $S$, which describes a condition of the experiment.

Let's ground these definitions in an example. Consider an experiment wherein you roll a fair single six-sided dice. The sample space $S$ is the collection of all possible outcomes, so $S =$

$\{1,2,3,4,5,6\}$. Since the dice is fair, the probability of any element in $S$ appearing is $1/6$, so

$$p(1) = \frac{1}{6}, \quad p(2) = \frac{1}{6}, \quad p(3) = \frac{1}{6}, \quad p(4) = \frac{1}{6}, \quad p(5) = \frac{1}{6}, \quad p(6) = \frac{1}{6}.$$

It's not too hard to see that the sum of all these probabilities is 1. An event in this sample space might be $E = \{2,4,6\}$, which describes the outcome where the dice is even.

A more complicated experiment is to flip a fair coin three times. If H indicates heads, and T tails, the sample space of this experiment is

$$S = \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{HTT}, \mathrm{THH}, \mathrm{THT}, \mathrm{TTH}, \mathrm{TTT}\}.$$

The probability of any event is $1/8$. Let's say we want an event of the form *"at least two heads appears."* The subset corresponding to this event is

$$E = \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{THH}\}.$$

**Definition 4.8**

Two events $E$ and $F$ corresponding to the same sample space $S$ are said to be *mutually exclusive* if $E \cap F = \emptyset$.

Continuing with the example of flipping a coin three times, let $F$ be the event describing *"flipping at least two tails"*:

$$F = \{\mathrm{HTT}, \mathrm{THT}, \mathrm{TTH}, \mathrm{TTT}\}.$$

The events $E$ and $F$ are mutually exclusive since there's no overlap between them.

**Example 4.9**

Suppose an experiment is held where two fair six-sided die are thrown, and their faces are recorded. Determine the size of this sample space, and the probability that any single event will occur.

*Solution.* Using our knowledge of counting, the number of possible outcomes is the number of possibilities for the first dice (6) multiplied by the possibilities for the second dice (6), so there are 36 possible outcomes. Since the outcomes are equiprobable, the probability of any event happening is $1/36$. ∎

Given a sample space $S$, a probability distribution $p$, and an event $E$, the *probability that the event $E$ occurs* is

$$P(E) = \sum_{s_i \in E} p(s_i).$$

For example, if a fair coin is flipped three times and $E = \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{THH}\}$ is the event "at least two heads appear," then

$$P(E) = \sum_{s_i \in E} p(s_i) = p(\mathrm{HHH}) + p(\mathrm{HHT}) + p(\mathrm{HTH}) + p(\mathrm{THH}) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{1}{2}.$$

Note that in an equiprobable space $S$, the probability of each event is $p(s_i) = 1/|S|$. Thus if $E$ is an event, the probability of $E$ is

$$p(E) = \sum_{s_i \in E} p(s_i) = \sum_{s_i \in E} \frac{1}{|S|} = \underbrace{\frac{1}{|S|} + \frac{1}{|S|} + \cdots + \frac{1}{|S|}}_{|E|\text{-times}} = \frac{|E|}{|S|}.$$

**Example 4.10**

Consider an experiment where two fair six-sided die are thrown and their faces are recorded. Let $E$ be the event *"The first die thrown shows an even number."* Determine the probability of $E$.

*Solution.* We know the sample space $S$ consists of 36 equiprobable events. If we can determine the size of the event space, the solution will be the quotient of those two numbers. To determine the size of the event space we use the Basic Counting Principle. The first dice must be even, of which there are 3 possibilities. There is no restriction on the second dice, yielding 6 possibilities. Thus $|E| = 3 \times 6 = 18$, and the probability that the first dice is even is

$$P(E) = \frac{18}{36} = \frac{1}{2}.$$

■

**Example 4.11**

In the game *Settlers of Catan*, two dice are thrown and the sum of the shown numbers are used to determine production. If a 7 is rolled, the Robber comes into play. Assuming both dice are fair, what is the probability that a player invokes the Robber?

*Solution.* The sample space is the usual 36 element space derived from throwing two die, but the event needs more thought. One option is to write out the sum of all 36 combinations, but this is cumbersome. Instead, we think of all the ways a 7 could be rolled. This gives us

$$E = \{(1,6),\ (2,5),\ (3,4),\ (4,3),\ (5,2),\ (6,1)\}.$$

Thus the probability that the Robber comes into play is $P(E) = 6/36 = 1/6$. ■

**Example 4.12**

A fair coin is flipped six times. Determine the probability that exactly two heads appear.

*Solution.* This question is a bit trickier. The sample space $S$ consists of the set of all 6-tuples of H's and T's, and so has $2^6 = 64$ elements. Let $E$ be the event *"exactly two heads are flipped."* How do we count $E$?

The easiest way to solve this problem is to think of it a little differently. Consider the collection of numbers $\{1, 2, 3, 4, 5, 6\}$. We want to choose two numbers from this set, and these numbers will correspond to where the heads will occur in the flip. For example, choosing $(1,5)$ means that the flip looks like HTTTHT, while $(2,3)$ corresponds to THHTTT. Order does not matter, as $(1,5)$ and $(5,1)$ result in the same sequence of flips. Thus there are ${}_6C_2 = 6!/(4!2!) = 15$ possibilities for flipping exactly two heads. The probability of flipping exactly two heads is thus

$$P(E) = \frac{|E|}{|S|} = \frac{15}{64}. \qquad \blacksquare$$

What if the coin is not fair, but instead we're told that heads are twice as likely as tails. Because we are no longer in the equiprobable situation, our probability distribution will change. We need to figure out how to take a probability distribution on a single event, and turn it into a probability distribution on multiple events.

**Definition 4.13**

Two experiments are said to be *independent* if the outcome of one does not affect the other. If $S^A = \left\{s_1^A, \ldots, s_n^A\right\}$ and $p_A$ describe the sample space and probability distribution of the first experiment, and $S^B = \left\{s_1^B, \ldots, s_m^B\right\}$ and $p_B$ describe the sample space and probability distribution of the second experiment, then the probability of a joint event $s_i^A s_j^B$ is

$$P(s_i^A s_j^B) = p_A(s_i^A) p_B(s_j^B).$$

Suppose we are given a coin such that $p(\mathrm{H}) = 2/3$ while $p(\mathrm{T}) = 1/3$. We will flip the coin twice and record the result. These events are independent: It second flip of the coin does not depend on the result of the first flip. Therefore, the probability of the two-flip experiment is the product of the probabilities:

$$P(\mathrm{HH}) = \frac{2}{3} \times \frac{2}{3} = \frac{4}{9}, \quad P(\mathrm{HT}) = \frac{2}{3} \times \frac{1}{3} = \frac{2}{9}, \quad P(\mathrm{TH}) = \frac{1}{3} \times \frac{2}{3} = \frac{2}{9}, \quad P(\mathrm{TT}) = \frac{1}{3} \times \frac{1}{3} = \frac{1}{9}.$$

The sample space of the two-flip experiment is $S = \{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}$. If $E = \{HT, TH\}$ is the event *"exactly one tails is flipped,"*, we can use the probability distribution above to determine the probability of $E$,

$$P(E) = \sum_{s_i \in E} P(s_i) = P(\mathrm{HT}) + P(\mathrm{TH}) = \frac{2}{9} + \frac{2}{9} = \frac{4}{9}.$$

**Example 4.14**

Redo Example 4.10, but assume the first dice has a probability distribution

$$p(x) = \begin{cases} 7/12 & \text{if } x = 1 \\ 1/12 & \text{otherwise} \end{cases}.$$

*Solution.* The events are independent and the second dice is still equiprobable, so

$$P(1,n) = \frac{7}{12} \times \frac{1}{6} = \frac{7}{72} \quad \text{for} \quad n \in \{1, 2, \ldots, 6\}$$

and

$$P(k,n) = \frac{1}{12} \times \frac{1}{6} = \frac{1}{72} \quad \text{for} \quad k \neq 1 \quad \text{and} \quad n \in \{1,2,\ldots,6\}.$$

The event set $E$ still consists of 18 elements, and the first dice being even corresponds to $k = 2,4,6$, so

$$P(E) = \sum_{s_i \in E} p(s_i) = 18 \times \frac{1}{72} = \frac{18}{72} = \frac{1}{4}.$$ ■

If $E$ and $F$ are both events corresponding to a sample space $S$, the probability of $E$ *or* $F$ happening is represented by the union $E \cup F$. However, $E$ and $F$ could overlap, so when computing the probability of the union $E \cup F$, we have to ensure we don't count these twice. Thus

$$P(E \cup F) = P(E) + P(F) - P(E \cap F).$$

If $E$ and $F$ are mutually exclusive, then $E \cap F = \emptyset$ and this formula reduces to

$$P(E \cup F) = P(E) + P(F).$$

This tells us something convenient: If $E \subseteq S$ is an event, it is mutually exclusive to its complement $E^c$; that is, $E \cap E^c = \emptyset$. On the other hand $S = E \cup E^c$, giving

$$1 = P(S) = P(E \cup E^c) = P(E) + P(E^c).$$

We can rearrange this to read $P(E^c) = 1 - P(E)$.

While we're at it, if $E$ and $F$ are independent events, then $P(E \cap F) = P(E)P(F)$.

**Example 4.15**

Suppose a fair coin is flipped five times. What is the probability that at least one heads is flipped?

*Solution.* Our sample space $S$ consists of $2^5 = 32$ elements. Let $E$ correspond to the event *"at least one heads is flipped"*. Writing out $E$, or even counting the number of elements of $E$, is a non-trivial amount of work. Instead, note the $E^c$ is the event *"no heads are flipped."*, of which there is a single event $E^c = \{\text{TTTTT}\}$. Since the sample space is equiprobable,

$$P(E^c) = \frac{|E^c|}{|S|} = \frac{1}{32} \quad \text{so} \quad P(E) = 1 - P(E^c) = 1 - \frac{1}{32} = \frac{31}{32}.$$ ■

**Example 4.16**

Suppose 5 fair die are rolled simultaneously? What is the probability that at least two of the dice show the same number?

*Solution.* The sample space $S$ it the collection of all possible rolls of the five die, and so consists of $6^5 = 7776$ elements. Let $E$ be the event *"At least two of the dice show the same number."* It's difficult to count the elements of $E$, so again we look at $E^c$, which is *"Every dice shows a different* *number."* This isn't too bad: Order matters and there's no repetition, so this is a permutation. Hence $|E^c| = {}_6P_5 = 6!$, and

$$P(E^c) = \frac{6!}{7776} = \frac{5}{54} \quad \text{so} \quad P(E) = 1 - P(E^c) = \frac{49}{54} \approx 90.74. \qquad \blacksquare$$

On occasion, we need to find the probability of an event, but need to gather that data from a collection. For example, suppose a MAT133 classroom is sampled according to hair colour and height, resulting in the following table:

| | **Hair Colour** | | |
|---|---|---|---|
| **Height** | Blonde | Brunette | Redhead |
| Under 165cm | 9% | 10% | 4% |
| Between 165cm and 185cm | 21% | 29% | 16% |
| Above 185cm | 4% | 5% | 2% |

Suppose you want to find the probability a class member is a redhead. No individual element of the table will give you this result. However, every redhead in the class is either under 165cm, between 165cm and 185cm, or above 185cm, so by summing the probabilities of these events, you can effectively count all the redheads:

$$\begin{aligned} P(\text{redhead}) &= P(\text{redhead and height} < 165) + P(\text{redhead and } 165 \leq \text{height} \leq 185) \\ &\qquad + P(\text{redhead and height} > 185) \\ &= 0.04 + 0.16 + 0.02 = 0.22. \end{aligned}$$

**Law of Total Probability:** Let $S$ be the sample space of an experiment, with $F_1, F_2, \ldots, F_m$ a series of events forming a partition of $S$; that is, $F_1 \cup F_2 \cup \cdots \cup F_m = S$ and $F_i \cap F_j = \emptyset$ for all $i \neq j$. If $E \subseteq S$ is any other event, then

$$P(E) = P(E \cap F_1) + P(E \cap F_2) + \cdots + P(E \cap F_m) = \sum_{k=1}^{m} P(E \cap F_k).$$

## 4.3 Conditional Probability

The probability of an event happening might change if something about the state is already known. For example, the probability of developing lung cancer is known to be greater given that you are a smoker, compared to non-smokers. We have to find a way to build this additional information into our analysis. This being said, not all information is useful. For example, the probability of rolling a 6 on a fair dice given that you just flipped heads on a fair coin – the result of the coin makes no difference to the probability of the dice.

**Definition 4.17**

Let $S$ be the sample space for an experiment, with probability distribution $p$. Suppose $E, F \subseteq S$ are events. The *probability of $E$ given $F$*, written $P(E|F)$, is

$$P(E|F) = \frac{P(E \cap F)}{P(F)}. \tag{4.1}$$

We can motivate this formula as follows: Suppose the setup is the same as the definition, but that the distribution is equiprobable across all outcomes. Given that $F$ has already happened, we can throw away any event that is not in $F$ already; that is, we can restrict our sample space from $S$ to $F$. The probability of an event $E$ occurring, as restricted to the sample space $F$, is $E \cap F$, and so

$$P(E|F) = \frac{|E \cap F|}{|F|}. \tag{4.2}$$

To get to back to (4.1), we need to write these in terms of the original sample space $S$. This is done by multiplying and dividing by $|S|$:

$$P(E|F) = \underbrace{\frac{|E \cap F|}{|S|}}_{P(E \cap F)} \underbrace{\frac{|S|}{|F|}}_{P(F)^{-1}} = \frac{P(E \cap F)}{P(F)}.$$

This equation should then still hold when $p$ does not give an equiprobable distribution.

**Example 4.18**

A fair coin is flipped twice. What is the probability that both flips show heads, given that one of the flips is a heads?

*Solution.* The sample space is $S = \{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}$, and we define $E$ as the event where both coins are heads, and $F$ to be the event where at least one of the flips is a heads:

$$E = \{\mathrm{HH}\} \quad \text{and} \quad F = \{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}\}.$$

From here, $E \cap F = \{HH\}$, so $P(E \cap F) = 1/4$ and $P(F) = 3/4$, so

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

As $S$ consists of equiprobable events, we also could have used (4.2) to arrive at the same result. ∎

This might seen unintuitive – it probably seems as though the probability should be greater than 1/3. The important point here is that you're not told which flip admitted a heads, and this is what causes the problem. If the problem instead had said:

> "A fair coin is flipped twice. What is the probability that both flips are heads, given that *the first coin flip is heads*,"

we get a different result. Indeed, $E = \{\text{HH}\}$ and $F = \{\text{HH}, \text{HT}\}$, so $P(E|F) = 1/2$. The additional restriction that it was the *first* coin that flipped heads made all the difference.

**Example 4.19**

The Price-to-Earnings ratio (P/E) of a company is the ratio given by dividing the price of its stock with its earning per share. You've analyzed stocks trading on the University of Toronto Stock Exchange `UTSX`, and found the average P/E is 14. In addition, over the previous year, you found the following probabilities:

| | **P/E relative to average** | | |
|---|---|---|---|
| **Performance relative to average** | Low | Average | High |
| Underperforming | 3% | 11% | 9% |
| Average Performing | 10% | 19% | 5% |
| Outperforming | 25% | 13% | 5% |

For example, the probability that a company underperforms and has a high P/E ratio is 9%. Determine the probability that a company outperforms the market given that it has a high P/E ratio.

*Solution.* Let $E$ be the event *"A company has a high P/E ratio,"* and $F$ be the event *"A company outperforms the market average."* We're looking for $P(F|E)$, which we know can be evaluated as $P(F \cap E)/P(E)$. The probability $P(F \cap E) = 0.05$ can be read off from the table, so we need to find $P(E)$. Since under, average, and outperforming partition the market, the Law of Total Probability says that

$$P(E) = 0.09 + 0.05 + 0.05 = 0.19.$$

Thus the probability that a company outperforms the market, given it has a high P/E ratio, is

$$P(F|E) = \frac{P(F \cap E)}{P(E)} = \frac{0.05}{0.19} \approx 0.26. \qquad \blacksquare$$

If $E$ and $F$ are mutually exclusive, then $P(E \cap F) = \emptyset$, so $P(E|F) = 0$. Similarly, if $E$ and $F$ are independent then $P(E \cap F) = P(E)P(F)$, so

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{P(E)P(F)}{P(F)} = P(E).$$

This latter example is what we saw above, when two coins were flipped and we knew that the first coin was a heads.

We saw that (or can define) two independent events satisfy the relationship $P(E \cap F) = P(E)P(F)$, but until now we would have been forced to compute this by hand for non-independent events. Conditional probability gives us a new formula. If $E, F \subseteq S$ are two events, then

$$P(E \cap F) = P(E|F)P(F). \tag{4.3}$$

This has several advantages. The first is that we can rewrite the Law of Total Probability as follows: If $S$ is a sample space with events $F_1, \ldots, F_m$ forming a partition of $S$, and $E$ is some other event, then

$$P(E) = P(E \cap F_1) + P(E \cap F_2) + \cdots + P(E \cap F_m) = \sum_{k=1}^{m} P(E \cap F_k)$$
$$= P(E|F_1)P(F_1) + P(E|F_2)P(F_2) + \cdots + P(E|F_m)P(F_m) = \sum_{k=1}^{m} P(E|F_k)P(F_k).$$

**Example 4.20**

Suppose you are given two bags. The first bag consists of two white balls and one black ball. The second bag consists of two black balls and two white balls. Suppose a bag is selected at random, and a ball drawn from the bag. This ball is then placed into the other bag, and a new ball is chosen from that bag. What is the probability the ball is white?

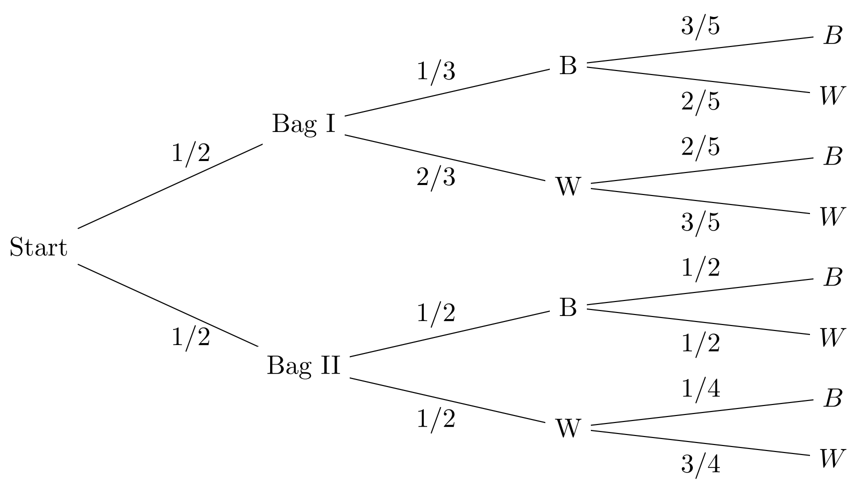



Figure 4.2: A probability tree describing conditional probabilities. We can use this tree to determine the solution to Example 4.20.

*Solution.* Consider the probability tree given in Figure 4.2, which can be constructed by conducting the eight different possible outcomes. This tree describes conditional probabilities. For example, in the second level, we have the probabilities

$$P(\text{B}|\text{Bag I}) = \frac{1}{3}, \quad P(\text{W}|\text{Bag I}) = \frac{2}{3}, \quad P(\text{B}|\text{Bag II}) = \frac{1}{2}, \quad P(\text{W}|\text{Bag I}) = \frac{1}{2}.$$

Using (4.3), we can compute the probabilities

$$P(\text{B} \cap \text{Bag I}) = P(\text{B}|\text{Bag I})P(\text{Bag I}) = \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}.$$

Graphically, this amounts to multiplying the products down the branches of the tree. Without going through all the computations explicitly, the second level probabilities are

$$P(\mathrm{B} \cap \text{Bag I}) = \frac{1}{6} \qquad P(\mathrm{W} \cap \text{Bag I}) = \frac{1}{3}$$
$$P(\mathrm{B} \cap \text{Bag II}) = \frac{1}{4} \qquad P(\mathrm{B} \cap \text{Bag II}) = \frac{1}{4}.$$

Similarly, the third level of the tree consists of more conditional probabilities, such as

$$P(\mathrm{B}|\mathrm{B} \cap \text{Bag I}) = \frac{3}{2} \quad \text{and} \quad P(\mathrm{B}|\mathrm{W} \cap \text{Bag II}) = \frac{1}{4},$$

though I won't write them all out. We're interested in determining when a black ball is drawn last, meaning all the paths which end in a black ball. We then sum over all of these paths. Doing this we get

$$\begin{aligned} P(B \text{ second draw}) &= P(\mathrm{B}|\mathrm{W} \cap \text{Bag I})P(\mathrm{W} \cap \text{Bag I}) + P(\mathrm{B}|\mathrm{B} \cap \text{Bag I})P(\mathrm{B} \cap \text{Bag I}) \\ &\quad + P(\mathrm{B}|\mathrm{W} \cap \text{Bag II})P(\mathrm{W} \cap \text{Bag II}) + P(\mathrm{B}|\mathrm{B} \cap \text{Bag II})P(\mathrm{B} \cap \text{Bag II}) \\ &= \left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{4}\right) + \left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\right) + \left(\frac{1}{2} \times \frac{2}{3} \times \frac{2}{5}\right) + \left(\frac{1}{2} \times \frac{1}{3} \times \frac{3}{5}\right) \\ &= \frac{1}{16} + \frac{1}{8} + \frac{2}{15} + \frac{1}{10} = \frac{101}{240}. \end{aligned}$$

■

**Theorem 4.21: Bayes' Theorem**

If $S$ is the sample space of some experiment with $E, F \subseteq S$ events in $S$, then

$$P(F|E) = \frac{P(E|F)P(F)}{P(E)}. \tag{4.4}$$

The derivation of this formula is straightforward: By (4.3) we know $P(E \cap F) = P(E|F)P(F)$, so

$$P(F|E) = \frac{P(F \cap E)}{P(E)} = \frac{P(E|F)P(F)}{P(E)}$$

which is Bayes' Theorem. Why do we care? In practice, we can measure $P(E|F)$ using prior data, but want to know $P(F|E)$ to make predictions about the future. For example, suppose we're trying to determine the probability that you pass MAT133 based off your Term Test 1 score. Let

$$\begin{aligned} T_A &= \text{"Scored above 60 on TT1"} \\ T_B &= \text{"Scored below 60 on TT1"} \\ S &= \text{"Passed MAT133"} \\ F &= \text{"Failed MAT133"} \end{aligned}$$

I can use prior years' data to determine $P(T_A|S)$; that is, the probability that you scored above a 60 on Term Test 1 given that you passed the course. Of course, as a student currently taking this course, you're more interested in the other conditional probability: $P(S|T_A)$ – the probability that you pass the course given that you score above a 60 on Term Test 1.

**Example 4.22**

From the above example, suppose there are 559 students, of which 325 scored above 60 on Term Test 1, 339 passed the course, and $P(T_A|S) = 0.776$. Determine $P(S|T_A)$.

*Solution.* Note that $P(S)/P(T_A) = 339/325 = 1.04$, since in either case we will divide by overall number of students. From (4.4), we know that

$$P(T_A|S) = \frac{P(S|T_A)P(T_A)}{P(S)} = 0.776 \times 1.04 = 0.809.$$

Hence you have an 81% probability of passing the course if you score above a 60 on Term Test 1. ■

## 4.4 Applications

With the tools of probability in hand, we can discuss some more advanced and interesting examples of probability.

### 4.4.1 Expected Value

The expected value is a way of a measuring the mean result of an experiment.

**Definition 4.23**

If $S$ is the sample space of some experiment, a *random variable* $X$ on $S$ is a real-valued function on $S$.

Random variables are a bit weird to think about, but the idea is that they are functions which depend on the experiment itself. For example, say our experiment is to flip a coin three times, so that the sample space is

$$S = \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{HTT}, \mathrm{THH}, \mathrm{THT}, \mathrm{TTH}, \mathrm{TTT}\}. \tag{4.5}$$

One choice of random variable could count the number of tails which appear, so that as a function:

$$X(s) = \begin{cases} 0 & s = \mathrm{HHH} \\ 1 & s = \mathrm{HHT},\ \mathrm{HTH},\ \mathrm{THH} \\ 2 & s = \mathrm{HTT},\ \mathrm{THT},\ \mathrm{TTH} \\ 3 & s = \mathrm{TTT} \end{cases}. \tag{4.6}$$

If $S$ has a probability distribution $p$, then we can make sense of the statement $P(X = 2)$, which reads "The probability that two tails are flipped". We can also write statements such as $P(X \geq 1)$ for "At least one tails is flipped," or $P(X \in \{1, 3\})$ for "Either 1 or 3 tails are flipped."

**Definition 4.24**

Suppose $S$ describes the sample space of an experiment, with probability distribution $p$. If $X$ is a random variable on $S$ with range $R$, the *expected value of $X$* is

$$\mathbb{E}[X] = \sum_{r \in R} r P(X = r). \tag{4.7}$$

The expected value is the mean value of the $X$; that is, it describes the average value of $X$ if the experiment were to be performed a large number of times.

**Example 4.25**

Let a fair coin be tossed three times, and let $X$ be the random variable describing the number of tails that are flipped. Determine the expected value of $X$.

*Solution.* The sample space $S$ is described in (4.5), and since the coin is fair we know that the probability distribution is equiprobable. The values that $X$ can take on are listed in (4.6), and consist of $R = \{0, 1, 2, 3\}$. To evaluate the expected value, (4.7) says we need to find the probabilities $P(X = r)$, from which we have

$$P(X = 0) = \frac{1}{8}, \quad P(X = 1) = \frac{3}{8}, \quad P(X = 2) = \frac{3}{8}, \quad P(X = 3) = \frac{1}{8}.$$

The expected value of $X$ is thus

$$\mathbb{E}[X] = \sum_{r \in R} r P(X = r) = \left[0 \times \frac{1}{8}\right] + \left[1 \times \frac{3}{8}\right] + \left[2 \times \frac{3}{8}\right] + \left[3 \times \frac{1}{8}\right] = \frac{12}{8} = 1.5.$$

Thus if we performed this experiment many times, we would expect to the average number of tails to be 1.5. ■

**Example 4.26**

A roulette wheel consists of 18 red spaces, 18 black spaces, and 2 green spaces. If you bet \$1 on red and the ball lands in a red space, you win an addition \$1 and lose your money otherwise. Let $X$ be the random variable which describes your winnings on a \$1 red bet. What is the expected value of $X$?

*Solution.* The probability of landing in a red space is $18/38$, while landing in a non-red space is $20/38$. The sample space for the experiment is $S = \{\text{red}, \text{black}, \text{green}\}$, and the value of the random variable is

$$X(s) = \begin{cases} 1 & \text{if } s = \text{red} \\ -1 & \text{if } s = \text{black} \\ -1 & \text{if } s = \text{green} \end{cases}.$$

Thus the expected value of $X$ is

$$\mathbb{E}[X] = \left[1 \times \frac{18}{38}\right] + \left[-1 \times \frac{18}{38}\right] + \left[-1 \times \frac{2}{38}\right] = -0.053.$$

This means that if you played roulette over the long term, consistently betting \$1 on red, you would walk away with \$0.95 at the end of the night. ■

The computation in Example 4.26 is important for the gambling industry. For the house to profit, it's necessary that the players always lose on average. On the other hand, you don't want the odds to be overwhelmingly against the players, otherwise they won't play. By skimming a small margin on average, Casinos can make healthy profits.

**Example 4.27**

You're holding a portfolio worth \$100,000 with Canadian bonds at 40% and international equity at 60%. You predict that bonds will change in price by 5% either way by the end of the year, with a 60% probability of increasing. Similarly, your equity holdings will change by 20% either way by the end of the year, with a 55% chance of increasing. What is the expected value of your portfolio in one year?

*Solution.* Let $X_B$ and $X_E$ be the random variables describing the value of your bond and equity holdings. You have \$40,000 in bonds, which will either increase to \$42,000 or decrease to \$38,000. Your \$60,000 in equity will change to either \$72,000 or \$48,000. Thus

$$\begin{aligned} \mathbb{E}[X_B] &= 42000 \times 0.6 + 38000 \times 0.4 = \$40,400 \\ \mathbb{E}[X_E] &= 72000 \times 0.55 + 48000 \times 0.45 = \$61,200 \end{aligned}$$

meaning the expected value of your portfolio is \$101,600 by the end of the year. ■

### 4.4.2 Markov Chains

Markov chains are used to model probabilistic systems whose evolution only depends on the previous state of the system. Integral to the study of a Markov chain is a matrix consisting of conditional probabilities, which describe the likelihood of transitioning from one state to another.

For example, suppose you're trying to model unemployment in the province. Using historical data over one year periods, you are able to determine the probability that a person is (un)employed given that that they were (un)employed in the previous year. Let $E_n$ and $E_p$ mean "employed now" and "employed previously" respectively, while $U_n$ and $U_p$ mean unemployed now and unemployed previously, respectively. You find that the probabilities are

$$\begin{array}{ll} P(U_n|U_p) = 0.4 & \qquad P(U_n|E_p) = 0.1 \\ P(E_n|U_p) = 0.6 & \qquad P(E_n|E_p) = 0.9. \end{array}$$

If we know that last year 92% of the population was employed and 8% was unemployed, then next year the percentage of employed and unemployed people will be

$$\begin{aligned} P(U_N) &= P(U_N|U_p)0.08 + P(U_n|E_p)0.92 = 0.124 \\ P(E_N) &= P(E_N|U_p)0.08 + P(E_n|E_p)0.92 = 0.876 \end{aligned}$$

This information can be summarized using matrices! Define the stochastic matrix $P$ and initial state $\mathbf{s}$ as

$$P = \begin{array}{c c} & \begin{array}{cc} U & E \end{array} \\ \begin{array}{c} U \\ E \end{array} & \begin{pmatrix} P(U_n|U_p) & P(U_n|E_p) \\ P(E_n|U_p) & P(E_n|E_p) \end{pmatrix} \end{array} = \begin{array}{c c} & \begin{array}{cc} U & E \end{array} \\ \begin{array}{c} U \\ E \end{array} & \begin{pmatrix} 0.4 & 0.1 \\ 0.6 & 0.9 \end{pmatrix} \end{array}, \qquad \mathbf{s} = \begin{bmatrix} 0.08 \\ 0.92 \end{bmatrix}.$$

Notice that the columns of $P$ and $\mathbf{s}$ sum to 1, since we're working with probabilities. To determine the probability of transitioning from one state to another over the period of a year, we can just compute $P\mathbf{s}$:

$$P\mathbf{s} = \begin{bmatrix} 0.4 & 0.1 \\ 0.6 & 0.9 \end{bmatrix} \begin{bmatrix} 0.08 \\ 0.92 \end{bmatrix} = \begin{bmatrix} 0.124 \\ 0.876 \end{bmatrix}.$$

The year after that can be computed by again multiplying by $P$,

$$P^2\mathbf{s} = P(P\mathbf{s}) = \begin{bmatrix} 0.4 & 0.1 \\ 0.6 & 0.9 \end{bmatrix} \begin{bmatrix} 0.124 \\ 0.876 \end{bmatrix} = \begin{bmatrix} 0.1372 \\ 0.8628 \end{bmatrix}.$$

If we compute the (un)employment after a few more years, we get

$$P^3\mathbf{s} = \begin{bmatrix} 0.141 \\ 0.859 \end{bmatrix}, \quad P^4\mathbf{s} = \begin{bmatrix} 0.142 \\ 0.858 \end{bmatrix}, \quad P^5\mathbf{s} = \begin{bmatrix} 0.143 \\ 0.857 \end{bmatrix},$$

where it appears as though the system is stabilizing to a single value. This is known as the *steady-state* for the system described by $P$.

**Theorem 4.28**

If $P$ is an $n \times n$ stochastic matrix (sum of its columns are 1) and $\mathbf{s}_0$ describes the probability of being one of those $n$-states , then

1. The system has a steady state solution; namely, $\lim_{n\to\infty} P^n \mathbf{s}_0$ exists,
2. The steady state solution $\mathbf{s}$ is the $\lambda = 1$ eigenvector of $P$; that is, $P\mathbf{s} = \mathbf{s}$.

**Example 4.29**

Standard and Poor regularly release the likelihood for a bond to transition between various credit ratings in a year. We'll use the following simplified model, which describes the probability transition matrix for a bond to change credit ratings:

$$P = \begin{array}{c c} & \begin{array}{ccc} AAA & BBB & CCC \end{array} \\ \begin{array}{c} AAA \\ BBB \\ CCC \end{array} & \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.7 & 0.2 \\ 0.1 & 0.2 & 0.7 \end{pmatrix} \end{array}.$$

Suppose the distributions of bonds from AAA to CCC are currently given by $\begin{bmatrix} 0.75, 0.15, 0.1 \end{bmatrix}$. Determine the steady-state bond distribution.

*Solution.* We row reduce the matrix $P-I$ to get

$$P - I = \begin{bmatrix} -0.2 & 0.1 & 0.1 \\ 0.1 & -0.3 & 0.2 \\ 0.1 & 0.2 & -0.3 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 0.1 & -0.3 & 0.2 \\ -0.2 & 0.1 & 0.1 \\ 0.1 & 0.2 & -0.3 \end{bmatrix}$$
$$\xrightarrow[-R_1+R_3\to R_3]{2R_1+R_2\to R_2} \begin{bmatrix} 0.1 & -0.3 & 0.2 \\ 0 & -0.5 & 0.5 \\ 0 & 0.5 & -0.5 \end{bmatrix} \xrightarrow[-2R_1\to R_1]{R_3+R_2\to R_2} \begin{bmatrix} 0.1 & -0.3 & 0.2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$
$$\xrightarrow{10R_1\to R_1} \begin{bmatrix} 1 & -3 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{3R_2+R_1\to R_1} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

This gives a one-parameter family of solutions $\begin{bmatrix} t, t, t \end{bmatrix}$ for any real number $t$. The additional requirement that the sum of the entries of this vector have to be 1 gives $3t = 1$ or $t = 1/3$. Thus the steady state distribution is

$$\begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

Or an equal amount of AAA, BBB, and CCC bonds. ■

### 4.4.3 Binomial Evolution

Suppose an experiment with two outcomes $\{s_0, s_1\}$ is performed – such as tossing a coin – and outcome $s_0$ occurs with probability $p$, so that $s_1$ must occur with probability $1-p$. If this experiment is repeated $n$-times, we're looking for a quick and easy way of describing the probability that $s_0$ appears $k$-times. We've seen this kind of argument before, but are now going to generalize it.

Let's start with an example. Suppose $k = 2$ and $n = 5$, and let's write down a representative possibility 00111, where a 0 indicates that $s_0$ occurred, and a 1 indicates that $s_1$ occurred. The probability of this single event is $p^2(1-p)^3$, but it is hardly the only event with exactly two 0's. To count the number of outcomes with exactly two zeroes, we think of the problem as saying "We have five places that event $s_0$ could occur, and we have to choose 2 of them," with order not mattering. The number of such outcomes is thus ${}_5C_2$. Thus the probability that $s_0$ will occur exactly twice in 5 trials is ${}_5C_2 p^2(1-p)^3$.

**The Binomial Distribution:** Suppose an experiment has two outcomes, $\{s_0, s_1\}$, with the probability of $s_0$ being $p \in [0,1]$, and the probability of $s_1$ being $1-p$. If the experiment is repeated $n$-times, the probability that $k$ of the trials result in $s_0$ is

$$B(n,k) = {}_nC_k p^k (1-p)^{n-k}.$$

The binomial distribution is valuable when modelling several periods of time where only one of two outcomes is likely, or modelling situations that can be coarsely modelled as such. For example, we might model stock projections using a binomial distribution, arguing that the stock will either increase or decrease in price by a projected amount each month, and aim to model the outcomes after a year. Similarly, increases and decreases to the prime interest rate can be modelled with a binomial distribution. We'll see an in-depth use of binomial pricing in Section 4.4.4 to determine the value of a option.

**Example 4.30**

The stock `UTM` is currently trading at \$10.00/share, with analysts projecting a 10% increase month-over-month with probability 60%, and a 10% decrease month-over-month with probability 40%. Determine the possible stock prices after 3 months, and the probabilities of each price. If $X$ is a random variable describing the stock price, determine the expected value of $X$ after those three months.

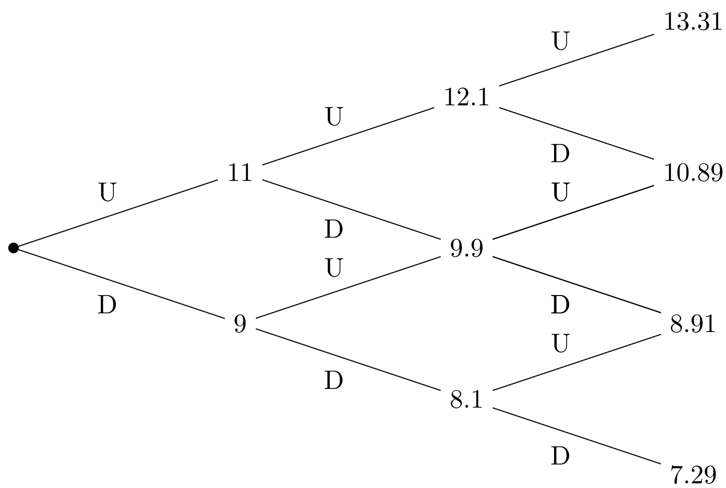



Figure 4.3: The collection of all possible movements in the stock price corresponding to `UTM`.

*Solution.* Matters are simplified by the fact that the order of a 10% increase or 10% decrease does not matter. Let U denote an increase in stock price, and D denote a decrease, in which case there are four possible outcomes as illustrated in Figure 4.3. Up to possible reordering, there are four possible outcomes for the stock with corresponding prices

$$\begin{aligned} &\text{UUU} : 10 \times 1.1^3 = \$13.31 && \text{UUD} : 10 \times 1.1^2 \times 0.9 = \$10.89 \\ &\text{UDD} : 10 \times 1.1 \times 0.9^2 = \$8.91 && \text{DDD} : 10 \times 0.9^3 = \$7.29. \end{aligned}$$

The probability that the stock goes up is $p = 0.6$. Let $B(3, k)$ denote the probability of the stock going up $k$-times, so that

$$\begin{aligned} &\text{UUU} : B(3,3) = ({}_3C_3)0.6^3(1-0.6)^0 = 0.216 && \text{UUD} : B(3,2) = ({}_3C_2)0.6^2(1-0.6)^1 = 0.432 \\ &\text{UDD} : B(3,1) = ({}_3C_1)0.6^1(1-0.6)^2 = 0.288 && \text{DDD} : B(3,0) = ({}_3C_0)0.6^0(1-0.6)^3 = 0.064. \end{aligned}$$

The expected value of the stock is thus

$$\begin{aligned} \mathbb{E}[X] &= (13.31 \times 0.216) + (10.89 \times 4.32) + (8.91 \times 0.288) + (7.29 \times 0.064) \\ &= \$10.61. \end{aligned}$$ ■

### 4.4.4 Options Pricing

An *option* is a derivative which allows its buyer to buy or sell a particular security at a fixed, pre-determined price. An option in which you can buy a security is a *call option*, while an option which allows you to sell a security is a *put option*. The price at which you buy or sell the security is known as the *strike price*. Options are further categorized into European and American options: European options can only be exercised at the expiration of the contract, while American options can be exercised at any time during the contract.

For example, say `UTM` is a stock currently trading at \$20.00 per share. You purchase a strike-25 European call option which expires in 6 months. This means that in 6 months when the option expires, you can purchase shares of `UTM` for \$25.00 each. If `UTM` is trading at \$30/share in 6 months, your options are worth \$5.00, since you can purchase `UTM` for \$5 cheaper than its current trading value. If `UTM` is trading for less than \$25 per share, your options are worthless. Note that, as the name indicates, you are not obligated to exercise the contract (in contrast to another derivative called *futures*).

One of the famous methods for pricing options is the Black-Scholes equation – a stochastic differential equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0,$$

which describes the price $V$ of an option as a function of stock price $S$ and time $t$. Here $r$ and $\sigma$ are constants. Getting to the point where one can solve this requires a great deal of mathematics. Instead, we can use a binomial pricing model to facilitate options pricing.

To price an option, we create a portfolio which replicates the returns on the option. This is called a *replicating portfolio*. Let $\Delta$ be the number of shares of `UTM` you buy, at a share price $S$. Suppose you also lend $B$ dollars (so if $B$ is negative you're borrowing) at a risk free rate of $r\%$ annually. Assume shares of `UTM` pay no dividend, and that the interest on $B$ compounds continuously.

Your replicating portfolio consists of the shares of `UTM` and the value of the amount lent/borrowed. At time $t = 0$, the value $C_0$ of the account is the equity value $\Delta S$ plus the money in the account $B$. Hence $C_0 = \Delta S + B$. After $h$-years, we assume the value of the stock either goes up by a ratio of $u$, or down by a ratio of $d$. For example, if $S = \$20$, $u = 1.3$ and $d = 0.8$, then after $h$ years the stock is either worth $uS = \$26$ or $dS = \$16$. The two possible values $C_h$ of the replicating portfolio after $h$-years are

$$C_h = \begin{cases} \Delta uS + Be^{rh} & \text{if the stock goes up} \\ \Delta dS + Be^{rh} & \text{if the stock goes down.} \end{cases}$$

Let $C_u$ be the price of the option after $h$ years assuming the price of the stock goes up. Similarly, let $C_d$ be the value if the price of the stock goes down. The value of the replicating portfolio must coincide with these prices, giving the linear system

$$\begin{aligned} \Delta uS + Be^{rh} &= C_u \\ \Delta dS + Be^{rh} &= C_d \end{aligned}$$

We can solve this system for the number of shares $\Delta$ and the lending value $B$ in terms of $C_u, C_d, S, r$, and $h$, to get

$$\Delta = \frac{C_u - C_d}{S(u-d)}$$
$$B = e^{-rh}\frac{uC_d - dC_u}{u-d}.$$

Substituting this into $C_0$ gives

$$C_0 = \Delta S + B = \frac{C_u - C_d}{u-d} + e^{-rh}\frac{uC_d - dC_u}{u-d}$$

**Example 4.31**

Suppose you buy a strike-22 European call option on `UTM`, currently trading at \$20, and set to expire in 1 year. Determine the price of that option if $u = 1.3, d = 0.8$ with a risk free interest rate of 4%.

*Solution.* As mentioned above, a strike-22 call option has values $C_u = \$4$ and $C_d = \$0$. Substituting everything into $C_0$ we get

$$C_0 = \frac{4-0}{1.3-0.8} + e^{-0.04}\frac{(1.3\times 0) - (0.8\times 4)}{1.3-0.8} = \$1.85. \qquad \blacksquare$$

## 4.5 Exercises

4-1. In a standard game of Yahtzee you roll 5 dice, aiming for certain combinations. How many possible outcomes are there for your first roll?

4-2. You play the following terrible game: Flip a coin, roll a die, and pick a card from a deck. How many different outcomes are there to the game?

4-3. You are in an escape room with four of your friends. There are five light switches on a wall. You must have each light switch in the correct "ON, OFF" setting in order to unlock a trunk. You have discovered the clues for Switch 1 and Switch 4, but cannot find the clues for the other three switches. Time is running out, and you decide to brute force it the problem. What is the maximum number of configurations you must try?

4-4. Consider a standard deck of 52 cards, from which 5 cards will be dealt.

(a) A one-pair occurs when you have two cards of equal rank, and the three remaining cards have a different, unequal rank. For example, 3♠ 3♣ 8♡ 10♢ J♣ is a one-pair. Determine the number of one-pair hands that can be dealt.

(b) A three-of-a-kind occurs when you have three cards of the same rank, with the other two cards having different, unequal rank. For example 5♡ 5♠ 5♢ 7♣ $Q$♠ is a three-of-a-kind. Determine the number of three-of-a-kind hands that can be dealt.

(c) mark) A full house is a three-of-a-kind of one rank, with the remaining two cards forming a one-pair. For example, 10♡ 10♢ 10♣ Q♣ Q♠. Determine the number of full-house hands one can be dealt.

4-5. A *straight* is a poker hand with 5 consecutively numbered cards, and suit does not matter. For example,

$$3\heartsuit, 4\diamondsuit, 5\heartsuit, 6\clubsuit, 7\clubsuit$$

is a straight. A *flush* is any five cards of the same suit, so

$$A\spadesuit, 6\spadesuit, 7\spadesuit, J\spadesuit, K\spadesuit$$

is a flush.

(a) In a standard deck of 52 cards, how many straights are there?

(b) In a standard deck of 52 cards, how many flushes are there?

(c) A *straight flush* is both a straight and a flush. How many straight flushes are there?

(d) A *royal flush* is a straight flush whose lowest rank card is a 10. How many royal flushes are there?

4-6. (a) How many arrangements of the word CHANCELLOR are there?

(b) How many arrangements of the word CHANCELLOR have all vowels in alphabetical order (but not necessarily beside each other)?

(c) How many arrangements of the word CHANCELLOR have all letters in alphabetical order?

4-7. Consider the word "FINANCE."

(a) How many six letter words can be formed by rearranging these letters?

(b) How many six letter words can be formed if the first three letters must be CAN?

(c) How many six letter words can be formed if the word must start with a vowel?

(d) How many four letter words can be formed if no vowels can be used?

(e) How many three letter words can be formed if no vowels can be used?

(f) How many words of any length can be formed if no vowels can be used?

(g) How many words of any length can be formed if consonants must alternate with vowels?

4-8. Now you are coaching a full hockey team, with 20 players. Once again, we'll assume every player can (somehow) play every position. You need four forward lines (each with three forwards), three defensive lines (each with two defensive players), and two goalies. How many different line combinations are there?

4-9. How many five digit numbers contain no repeated digits, have no even digits, and the sum of their digits is 25?

4-10. You are judging a dog show. There are seven contestants, and you must choose to rank the top three (the other four do not place). How many ways are there of doing this?

4-11. How many possible 5 card poker hands can be formed from a deck of 52 cards?

4-12. You're coaching a peewee hockey team. You have 13 non-goalie players, and need to assign them to one of the five positions for the starting lineup. How many different combinations of players do you have?

4-13. *Two-pair* is a poker hand consisting of two cards of the same rank, two cards of the same rank but different from the first, and a fifth card different from the rank of the first two. For example

$$3\heartsuit, 3\clubsuit, 10\heartsuit, 10\diamondsuit, A\clubsuit$$

is a two-pair hand. How many two-pair hands are there in a standard deck of 52 cards?

4-14. Consider a bag consisting of 10 marbles, of which 6 are red and 4 are blue.

(a) If you draw 5 marbles from the bag (without replacement), how many different samples could you draw?

(b) How many of those 5-marble samples consist of 3 red marbles?

(c) How many of those 5-marble samples consist of at least 3 red marbles?

4-15. Two dice – coloured red and blue – are thrown, and their values recorded. Consider the events:

- $E_1$ = Both dice show the same number,
- $E_2$ = Both dice are even,
- $E_3$ = Both dice are odd,
- $E_4$ = The red dice is strictly larger than the blue dice,
- $E_5$ = The sum of the dice is strictly greater than 5.

(a) Write out the sets corresponding to each event.

(b) Determine which of the events (if any) are mutually exclusive.

4-16. The following problem is somewhat long, but interesting. Consider two standard dice, with faces $\{1, 2, 3, 4, 5, 6\}$.

(a) When the two dice are rolled, the sum of their values can be anything between 2 and 12. Determine the number of ways of getting a sum of 2, a sum of 3, a sum of 4, etc.

(b) Consider two non-standard dice. The first has the faces $\{1, 2, 2, 3, 3, 4\}$ and the second has the faces $\{1, 3, 4, 5, 6, 8\}$. Once again the dice are summed, with the lowest possible value a 1, and the highest a 12. Determine the number of ways of getting a sum of 2, a sum of 3, a sum of 4, etc.

(c) Comparing your numbers from (a) and (b), conclude that playing with the non-standard dice give you exactly the same outcome as playing with the two standard dice.

This is the only pair of non-standard dice which give the same outcome as a standard pair.

4-17. Three fair coins are flipped and their values notes. Determine the probability of each event.

(a) All three flips are heads.

(b) The first flip is tails and the last is heads.

(c) At least two flips are heads.

(d) Exactly one flip is heads.

4-18. Consider the roll of a standard six-sided die, and choosing one card from a standard deck of 52 cards. Determine the probability of each event.

(a) The die shows an even number, and a red card is chosen.

(b) The die is a number strictly bigger than 4, and the card is an ace.

(c) The card has rank strictly larger than 10 (assume aces are low).

4-19. Determine the probability of being dealt each of the following hands from a single deal of a standard 52 card deck.

(a) Two pair

(b) Straight

(c) Flush

(d) Straight flush

4-20. You're playing a role playing game using the Generic Universal RolePlaying System (GURPS). The primary mechanism of GURPS is rolling skill checks. For example, suppose you have a Strength of 12 and want to move a large boulder. To pass your skill check, you roll three six-sided dice and sum their faces. If the total value you roll is *less than or equal* to 12, you pass your skill check and move the boulder. Otherwise, you fail.

You encounter an angry mob of goblins, and want to sneak by them. To do this, you must roll a skill check against your Stealth stat, which is 15.

(a) What is the probability that you successfully pass your skill check, and sneak by the goblins?

(b) Your skill check is a *critical success* if you roll either a 3 or a 4, or your margin of success (the skill level less your roll) is strictly greater than 10. For example, with a Perception skill of 20, you will critically succeed on any of a $3, 4, \ldots, 9$. In the example above where your Stealth skill is a 15, what is the probability of a critical success in your Stealth roll?

4-21. Suppose you have a blue die and a red die. The blue die is fair, but the red die has a probability distribution

$$r(x) = \begin{cases} 1/9 & \text{if } x \text{ is even} \\ 2/9 & \text{if } x \text{ is odd} \end{cases}.$$

Let $E$ be the event "One die is exactly twice the other" and let $F$ be the event "Both die are a multiple of 3."

(a) Write out $E$ and $F$.

(b) Are $E$ and $F$ mutually exclusive? Be sure to explain your answer.

(c) Determine $P(E)$ and $P(F)$.

4-22. A section of MAT133 was polled on their favourite sport and music, with the following results:

| | Hockey | Baseball | Basketball | Football | Hacky Sack |
|---|---|---|---|---|---|
| Rock | 22 | 10 | 12 | 4 | 0 |
| Pop | 15 | 8 | 15 | 8 | 2 |
| Hip Hop | 5 | 12 | 13 | 10 | 1 |
| Estonian Folk Metal | 2 | 1 | 1 | 0 | 4 |

(a) If a student is selected at random, what is the probability they like hockey?

(b) If a student is selected at random, what is the probability they like Estonian Folk Metal?

(c) If two students are chosen at random, what is the probability they both like basketball?

(d) If two students are chosen at random, what is the probability they both like the same music?

(e) If two students are chosen at random, what is the probability they both like hockey, or one of them likes Estonian Folk Metal?

4-23. Lenny is jumping hurdles. The probability that he jumps both hurdles is 75%, while the probability that he jumps the first hurdle is 90%. Having successfully jumped the first hurdle, what is the probability Lenny successfully jumps the second hurdle?

4-24. Carl draws two cards from a standard deck of 52 cards, without replacing the first.

4-25. Two fair dice are thrown, and their sum is 8. What is the probability that the first die is a 3?

4-26. Two fair dice are rolled.

(a) What is the probability both dice are even, given that the first dice is even?

(b) What is the probability both dice are even, given that one of the dice is even?

(c) What is the probability that both dice are even, given that the second dice shows a number strictly larger than 3?

(d) What is the probability that both dice are even, given that one of the dice shows a number strictly larger than 3?

4-27. Consider the following experiment:

- Roll a fair six-sided die (d6) once, and record the number as $X$,
- Then flip a fair coin $X$ times, and record the number of heads that appear as $Y$.

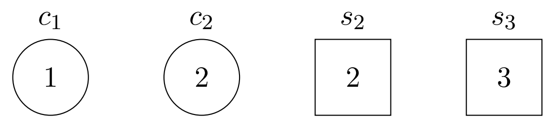
(a) Compute $P(Y = 4 \mid X = 6)$.

(b) Compute $P(Y = 4)$.

(c) Compute $P(X = 6 \mid Y = 4)$. *Hint:* Use your solutions to (i) and (ii).

4-28. You are given four shapes, each with a number inside of it. The first two shapes are circles with the numbers one (1) and (2), which we will denote by $c_1$ and $c_2$. The second two shapes are squares with the numbers two (2) and three (3), which we will denote by $s_2$ and $s_3$. You randomly select two shapes, all with equal probability, **and with repetition allowed**.

For example, $(c_1, c_1)$ means you chose $c_1$ both times, while $(c_2, s_3)$ means you first chose $c_2$ then $s_3$.




(a) How many ways are there to choose two shapes? Clearly explain your reasoning.

(b) Let $E$ be the event "Neither of the shapes are circles," and $F$ be the event "The sum of the chosen values is exactly four (4)." Write out the sets $E$ and $F$.

(c) Determine $P(E), P(F)$, and $P(F|E)$.

4-29. Consider the following experiment. You have two bags: Bag I contains 5 blue marbles and 2 red marbles, while Bag II contains 2 blue marbles and 3 red marbles. You begin by rolling two fair die. If the sum of the die total a number strictly greater than 10, you choose Bag I, and otherwise you choose Bag II. Now flip a fair coin. If the coin shows heads, add one blue marble to your bag, and add a red marble on tails. Now choose a marble at random from your bag. What is the probability that the chosen marble is red?

4-30. Suppose in a class of 560 students, the table below describes how those students performed on their first term test ($T$) and their final mark ($F$), as a percentage, in the course:

| | Final Score | | | | | |
|---|---|---|---|---|---|---|
| Test 1 | $F \in [0, 20)$ | $F \in [20, 40)$ | $F \in [40, 50)$ | $F \in [50, 60)$ | $F \in [60, 80)$ | $F \in [80, 100]$ |
| $T \in [0, 20)$ | 19 | 4 | 4 | 0 | 0 | 0 |
| $T \in [20, 40)$ | 17 | 35 | 8 | 6 | 5 | 0 |
| $T \in [40, 50)$ | 3 | 23 | 16 | 17 | 51 | 2 |
| $T \in [50, 60)$ | 2 | 24 | 9 | 21 | 2 | 2 |
| $T \in [60, 80)$ | 2 | 24 | 17 | 48 | 91 | 14 |
| $T \in [80, 100]$ | 0 | 3 | 1 | 7 | 36 | 47 |

We say that a student "Passed the term test" if they scored at least 50% in the term test, and "Passed the course" if they scored at least 50% on their final mark.

In each of the following questions, clearly explain how you arrived at your answer.

(a) What is the probability a student got 80%+ in the course?

(b) What is the probability a student failed the term test and passed (scored 50%+) the course?

(c) What is the probability a student passed the course, given that they passed the term test?

(d) What is the probability a student passed the course, given that they failed the term test?

(e) What is the probability a student got 80%+ in the course given that they passed the term test?

4-31. You are gambling on a chess tournament with three matches taking place:

| | | |
|---|---|---|
| Alexey | Oscar | Tyler |
| vs | vs | vs |
| Maha | Mahnoor | Parker |

All players are equally matched, so the probability of any player winning a match is $1/2$. You must guess which player will win each match. It costs you \$10 to play, with prizes as follows:

- Your original \$10 back, plus an additional \$20 if you guess all three matches correctly.
- Your original \$10 back if you guess exactly two of the matches correctly.
- An amount of \$2 if you guess a single match correctly.

(a) Let $X$ be the value of your winnings, so that $X = 0$ corresponds to breaking even, a negative value of $X$ corresponds to losing money, and a positive value of $X$ means you win money. What is the expected value of $X$?

(b) You bribe Tyler to deliberately lose his match, guaranteeing that Parker will win. Assuming you pick Parker to win, what is the expected value of your winnings now?

4-32. You and your friend Hossein play the following game: You each roll a six-sided dice. If Hossein's dice is larger, you pay him the difference of the shown values; if your dice is larger, Hossein pays you the difference; if your dice show the same amount, you both keep your money. For example, if $(1, 5)$ means you throw a 1, and Hossein throws a 5, then you must pay Hossein \$4. If the dice rolled are $(4, 3)$, then Hossein owes you \$1.

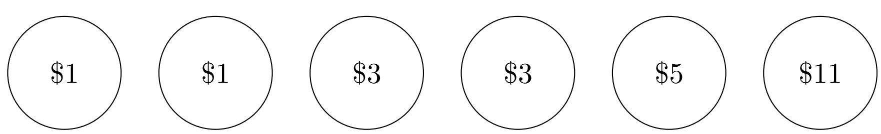
(a) (3 marks) What is the expected value of your winnings if you play this game? By winnings, we mean that the expected value is positive if you win money, negative if you lose money, and 0 if you break even.

(b) (2 marks) Now suppose you manage to acquire a die with the following probability distribution:

$$d(x) = \begin{cases} 1/3 & x = 4, 5, 6 \\ 0 & x = 1, 2, 3 \end{cases}.$$

What is the expected value of your winnings now?

4-33. At a carnival, the following game costs you \$5 to play. You reach into a basket which contains 6 balls, each with a number written on them. The draw is equiprobable, so you have an equally likely chance of picking any ball. The numbers are as follows:




You win the amount of money shown on the ball you have chosen. For example, if you select a ball labelled "\$3," you win \$3. In this case, your winnings are $\$3 - \$5 = -\$2$, where here a negative number means you lost money.

(a) What is the probability you choose a ball which results in you winning money?

(b) What is the probability you lose money given that you did not choose a ball labelled "\$1"?

(c) What is the expected value of your winnings? *Note:* This number should be negative if you lost money, and positive if you won money.

4-34. You are the chief of Toronto Police Services. There was a bank robbery downtown, and a witness has come to you stating that she saw the robber enter an orange taxi to flee the crime. You know that 60% of the taxis in the city are orange. You also know that witnesses are not completely reliable, and that a generic eye witness will correctly identify the color of the taxi only 90% of the time. Determine the probability that the criminal actually entered an orange taxi.

4-35. You are Vice President of Ticket Sales for Toronto FC. The average price of ticket is \$20, and you expect to sell 8000 tickets on any given day. However, if it rains you will only sell 1500 tickets. Tomorrow there is a 15% chance of showers. What is the expected sales revenue of tomorrow's game?

4-36. For this question, a d$n$ means a dice with $n$-sides whose faces are consecutively numbered from 1 to $n$. So for example, a d6 is a standard six-sided dice whose faces have the numbers $\{1, 2, 3, 4, 5, 6\}$, while a d4 is a four-sided dice with faces $\{1, 2, 3, 4\}$.

Figure 4.4: In order starting at the top left, a d4, d6, d8, d10, d12, d20.

You are given a bag consisting of a single 1d4, 3d6, and 1d20. You reach into this bag and draw out a dice at random, then roll that dice and record the number which appears. Assume that the probability you grab any given dice is equiprobable, and that all dice are fair. What is the expected value of your roll?

4-37. When a major book retailer purchases books, they can send any unsold books back to the publisher for a refund. You work for one of these major publishers, and are distributing your company's newest book. You calculate that 85% of your books in Toronto will sell, while 80% of the books in Mississauga will sell. Each book sold nets your company \$10 profit, while any unsold book nets a \$3 loss.

(a) Determine the expected return on a single book in Toronto.

(b) Determine the expected return on a single book in Mississauga.

(c) Determine the expected return on a single book overall, in both Toronto and Mississauga.

(d) If you ship 1000 books to Toronto, and 500 books to Mississauga, what is your expected profit?

4-38. Patients in the hospital undergo the following transitions on a day-to-day basis:

- Patients in surgery either go into intensive car or rehab, with equal probabilities, but never go back into surgery.
- Patients in intensive care or rehab have an even chance of staying in their intensive car or rehab.
- If a patient leaves intensive car or rehab, only half of the time do they re-enter surgery.

Patients in this hospital never get better – it's not a good hospital.

(a) Write down a matrix which describes the transition probabilities of each event.
(b) If a patient is in surgery today, what is the probability he is in surgery four days from now?
(c) Find the long term equilibrium behaviour of the hospital.

4-39. The Bank of Canada meets every six weeks to determine the change in the prime interest rate. Lubimir in Analysis has determined that over the next 24 weeks, any given meeting admits a 65% chance the central bank will raise the rates by 0.25%, and the rate will remain the same otherwise. If the current prime interest rate is 1.5%, what is the expected value of the prime interest rate in 24 weeks?

4-40. Repeat the options pricing derivation from Section 4.4.4, but now assume that the stock pays dividends. Assume the dividend is paid at a nominal rate of $\delta\%$ per annum, and is paid out continuously. Use the numbers in Example 4.31 to determine the option price.

# 5 Limits

## 5.1 Some Motivation

Limits are the method by which we, as manifestly finite beings, deal with concepts of infinities and infinitesimals. The goal towards which we are working is a description of instantaneous rate of change, so let's think on what this means.

The majority of us have been in a car at some point or another, and have afforded a casual glance at the speedometer. Let us say that at the instant we look down, the speedometer reads 90 km/hr. Have you ever thought about what it means, at that single instant in time, to be travelling at that speed? As suggested by its units, speed is an object which requires both distance and time to measure, but at a single moment, neither any time nor any distance has passed, so what does this mysterious quantity mean?

Despite my claims that the previous example should get you thinking about how the word "instantaneous" really affects a quantity, many of you will simply shrug aside my suggestions. In anticipation of this reaction, what if we change the associated quantities around and instead of the instantaneous speed of a car, we discuss shopping! At any given point of time, somebody on this planet is making a purchase. Assume that we were able to measure the rate at which people were spending money, and I told you that at this moment in time the human species was globally spending 140 million dollars an hour? What does this mean?

Now on the other hand, what if you were asked to determine the instantaneous speed of a race car at the instant its front bumper passes a finish line? Being clever students, you decide to measure how far the car has travelled in the minute before it hits the finish line, and get a result of 1500 meters. Hence the car was travelling

$$\frac{1500 \text{ metres}}{1 \text{ minute}} \times \frac{1 \text{ kilometre}}{1000 \text{ metres}} \times \frac{60 \text{ minutes}}{1 \text{ hour}} = \frac{90 \text{ kilometres}}{1 \text{ hour}}.$$

But what if the cars speed was not constant during that minute? What if the driver accelerated at the end? You decide that you can get a better estimate of the speed at the finish line by instead just looking at how far the car travelled in the single second before the car hit the finish line. This time the car travelled 30 metres, so you calculate

$$\frac{30 \text{ metres}}{1 \text{ second}} = \frac{108 \text{ kilometres}}{1 \text{ hour}}.$$

But still, this does not account for any change in acceleration which occurred in the last second. Your guess of 108 km/hr is probably close, but close is not good enough in mathematics! So you try again by measuring the distance after 0.1 seconds, then 0.01 seconds, and so on , but no matter how hard you try you cannot get the exact speed because there is always the chance that the car was not travelling at a constant speed during your measurements. Nonetheless, we know there must be an answer: the car was travelling at some speed, so what is it? Limits provide the solution.

### 5.1.1 Intuition

Limits are the mathematical device which allow us to infer information about a point by analyzing information about well-behaved points nearby. Let $f$ be an arbitrary function and $c \in \mathbb{R}$. We say that "the limit of $f(x)$ as $x$ approaches $c$ is equal to $L$" if, whenever we let $x$ get arbitrarily close to $c$ then $f(x)$ gets arbitrarily close to $L$. This is written as

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

The best way to gain an intuitive understanding of limits is to see a few examples. I would warn you that this first example is rather nicely behaved and fails to capture why we use limits. Nonetheless, simple examples are often the best for getting a grasp as to how something works.

**Example 5.1**

Consider the function $f(x) = 4x + 2$. Determine the limits

$$\lim_{x \to 0} f(x), \qquad \lim_{x \to -4} f(x), \qquad \lim_{x \to 5} f(x).$$

Form a hypothesis as to what the limit is as $x \to c$ for any value of $c$.

*Solution.* This solution is purely heuristic and is only presented in a way to show you how to think about these problems. The first example asks us to consider what happens when $x \to 0$, so we would like to see what happens for values of $x$ which are close (but not equal to zero). You can guess that as $x$ gets close to zero, $4x + 2$ gets close to $4 \cdot 0 + 2 = 2$. Similarly, as $x \to -4$ then

$4x+2$ approaches $4 \cdot (-4)+2 = -14$. The following table corroborates this idea:

| $x \to 0$ | | | | $x \to 0$ | | | |
|---|---|---|---|---|---|---|---|
| $x<0$ | | $x>0$ | | $x<0$ | | $x>0$ | |
| $x$ | $f(x)$ | $x$ | $f(x)$ | $x$ | $f(x)$ | $x$ | $f(x)$ |
| $-0.1$ | $1.6$ | $0.1$ | $2.4$ | $-4.1$ | $-14.4$ | $-3.9$ | $-13.6$ |
| $-0.05$ | $1.8$ | $0.05$ | $2.2$ | $-4.05$ | $-14.2$ | $-3.95$ | $-13.8$ |
| $-0.01$ | $1.96$ | $0.01$ | $2.04$ | $-4.01$ | $-14.04$ | $-3.99$ | $-13.96$ |
| $-0.005$ | $1.98$ | $0.005$ | $2.02$ | $-4.005$ | $-14.02$ | $-3.995$ | $-13.98$ |
| $-0.001$ | $1.9996$ | $0.001$ | $2.0004$ | $-4.001$ | $-14.004$ | $-3.999$ | $-13.996$ |
| $-0.0005$ | $1.9998$ | $0.0005$ | $2.0002$ | $-4.0005$ | $-14.002$ | $-3.9995$ | $-13.998$ |

As a matter of fact, it looks as though

$$\lim_{x\to 0} f(x) = f(0) = 2 \quad \text{and} \quad \lim_{x \to -4} f(x) = f(-4) = -14$$

so we guess that in general,

$$\lim_{x\to c} f(x) = f(c) = 4c+2.$$

■

In Example 5.1 we guessed that the limit as $x \to c$ could be determined by evaluating $f(c)$, and it turns out that in this example that is correct. However, we must be careful about just freely plugging in numbers into equations as the function might not always be defined at that point.

**Example 5.2**

Let $f(x) = \dfrac{x^2+x-6}{x-2}$. Determine the limit $\lim_{x\to 2} f(x)$.

*Solution.* Unlike the previous example, attempting the substitute $x=2$ into $f$ will result in division-by-zero, which we know is never permitted. However, we can evaluate $f$ at any number other than 2 and the hope is that this will tell us what the function looks like at $x=2$. Indeed,

| $x$ | $f(x)$ |
|---|---|
| 2.1 | 5.1 |
| 2.05 | 5.05 |
| 2.01 | 5.01 |
| 2.005 | 5.005 |
| 2.001 | 5.001 |
| 2.0005 | 5.0005 |

so it certainly appears as though $f$ is approaching 5. If $x \neq 2$ then we may factor $f$ as

$$\frac{x^2+x-6}{x-2} = \frac{(x+3)(x-2)}{x-2} = x+3$$

and the behaviour of this function as $x \to 2$ agrees with our observations. ■

The previous example demonstrates that a function does not need to be defined at a point for the limit at that point to exist. In fact, this is an excellent opportunity to point out that the functions $f(x) = (x^2+x-6)/(x-2)$ and $g(x) = x+3$ are similar but *are not equal*: the distinction being that the domain of $f(x)$ is $\mathbb{R} \setminus \{2\}$ while the domain of $g$ is $\mathbb{R}$. If two functions have different domains, they certainly cannot be equal! Of course, $x = 2$ is the only point where the functions do not agree, and their graphs are even identical with the exception that the graph of $f$ will have a hole at $x = 2$. This does not matter when we are taking limits, and we have the equality

$$\lim_{x\to 2} \frac{x^2+x-6}{x-2} = \lim_{x\to 2}[x+3].$$

While the functions differ at the point $x = 2$, the limit only looks at what the functions do at points close to *but not equal* to 2. Thus the limits see them as the same function (see Figure 5.1).

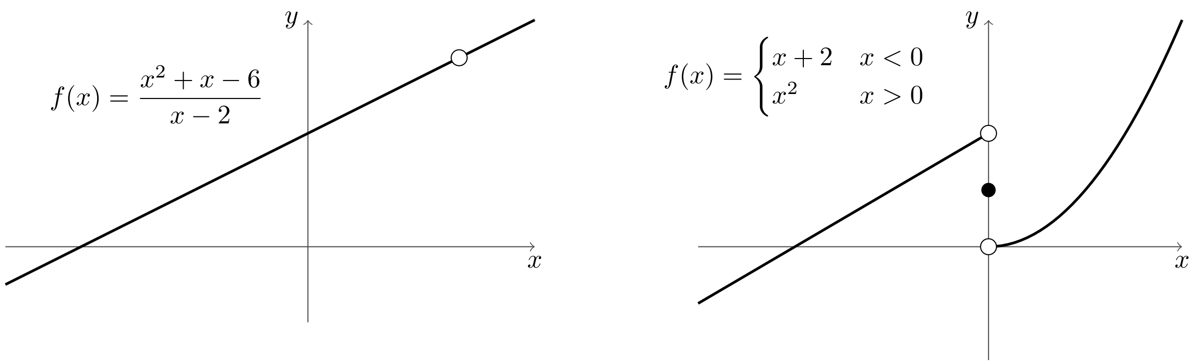



Figure 5.1: **Left:** The function $(x^2+x-6)/(x-2)$ is identical to the function $x+3$ except for the presence of a hole at $x = 2$. This does not affect the limit though, as the limit is only concerned with the behaviour of the function *near* $x = 2$. **Right:** A graph of a piecewise function whose limit at zero is dependent upon the direction of approach. Notice that in either case, the limit disagrees with the value of the function at zero.

**Example 5.3**

Compute the limit

$$\lim_{x\to 2} \frac{x-2}{\sqrt{x+7}-3}.$$

*Solution.* The usual thing to do in such situations where one summand contains a square root is to multiply by the conjugate. In this case, the conjugate is $\sqrt{x+7}+3$, and

$$\begin{aligned} \lim_{x\to 2} \frac{x-2}{\sqrt{x+7}-3}\frac{\sqrt{x+7}+3}{\sqrt{x+7}+3} &= \lim_{x\to 2} \frac{(x-2)(\sqrt{x+7}+3)}{(x+7)-9} \\ &= \lim_{x\to 2} \frac{x-2}{x-2}\left[\sqrt{x+7}+3\right] \\ &= \lim_{x\to 2}\left[\sqrt{x+7}+3\right] = 6. \end{aligned}$$

A shorter albeit more clever solution is to write

$$x - 2 = (x+7) - 9 = \left[\sqrt{x+7} - 3\right]\left[\sqrt{x+7}+3\right],$$

which leads to a cancellation argument similar to Example 5.2. ■

## 5.2 One Sided Limits

Implicit in our previous discussion of limits is that when we take $x \to c$, we must get the same answer whether we are approaching from the left of $c$ or the right of $c$. It is possible that approaching from the left and right actually give different values of the limit, as can be seen in Figure 5.1. This naturally leads to the idea of one-sided limits, where we restrict our attention to values of the function on only one side of the limiting point.

More generally, we will say that "the limit of $f(x)$ as $x$ approaches $c$ from the *right* is $L$" if whenever $x > c$ gets arbitrarily close to $c$, $f(x)$ gets arbitrarily close to $L$. This is written in symbols as

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

Similarly, we say that "the limit of $f(x)$ as $x$ approaches $c$ from the *left* is $L$" if whenever $x < c$ gets arbitrarily close to $c$, $f(x)$ gets arbitrarily close to $L$, and in this case we write

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

If both of the one-sided limits exist and are equal to the same value $L$, then the limit $x \to c$ exists and is also equal $L$. There are plenty of examples where the two-sided limit does not exist, as our following examples demonstrate.

**Example 5.4**

Consider the function

$$f(x) = \begin{cases} x+2 & x < 0 \\ 1 & x = 0 \\ x^2 & x > 0 \end{cases}.$$

Compute the limit of $f(x)$ as $x \to 0^-$ and as $x \to 0^+$. Does the two-sided limit exist?

*Solution.* We first look at the limit as $x \to 0^-$. In this case, we know that $x$ is always less than 0, so $f(x)$ effectively looks like the function $x + 2$. As $x \to 0^-$ we see that $x + 2 \to 2$ and so we conclude that

$$\lim_{x\to 0^-} f(x) = 2.$$

On the other hand, the limit $x \to 0^+$ guarantees that $x$ is always positive. Here, $f(x)$ looks like the function $x^2$ and as $x$ approaches 0, $x^2$ approaches 0 as well, so

$$\lim_{x\to 0^+} f(x) = 0.$$

Each one-sided limit exists, but they are not equal. Hence the two-sided limit does not exist. The graph of $f(x)$ is given in Figure 5.1. ■

The other way that a two-sided limit can fail to exist is if the one-sided limits do not exist either. This can happen in one of two ways: The first is that it is impossible to find a number $L$ to which the function gets close. This can happen, for example, if our function oscillates infinitely in any small interval around a point. Alternatively, the limits can *diverge*, meaning that the function goes to either positive or negative infinity. This next example demonstrates both phenomena.

## 5.3 Limit Laws

Mathematicians love to be lazy, in the sense that if we have already performed a calculation, why should we repeat it ever again? Similarly, we like to build complicated examples from simple examples. To this end, we formulate the following collection of limit laws, which are intended to dramatically simplify our life:

**Theorem 5.5: Limit Laws**

If $f$ and $g$ are functions such that

$$\lim_{x\to c} f(x) \quad \text{and} \quad \lim_{x\to c} g(x) \quad \text{both exist for some } c \in \mathbb{R}, \text{ then}$$

1. $\lim_{x\to c} [\alpha f(x)] = \alpha \lim_{x\to c} f(x)$,
2. $\lim_{x\to c} [f(x) \pm g(x)] = \left[\lim_{x\to c} f(x)\right] \pm \left[\lim_{x\to c} g(x)\right]$,
3. $\lim_{x\to c} [f(x)g(x)] = \left[\lim_{x\to c} f(x)\right] \left[\lim_{x\to c} g(x)\right]$,
4. $\lim_{x\to c} \left[\frac{f(x)}{g(x)}\right] = \frac{\lim_{x\to c} f(x)}{\lim_{x\to c} g(x)}$ provided $\lim_{x\to c} g(x) \neq 0$

**Example 5.6**

Compute the limit

$$\lim_{x\to 1} \frac{x^4 + 7x + 2}{x - 4}.$$

*Solution.* We would like to say that this is the limit of the quotient, and this will be the case so long as the limit of the numerator and denominator both exist, and the denominator is non-zero. For the numerator, we would again like to break this into a sum of limits:

$$\begin{aligned}
\lim_{x\to 1} \left[x^4 + 7x + 2\right] &= \lim_{x\to 1} x^4 + 7 \lim_{x\to 1} x + \lim_{x\to 1} 2 \\
&= \left[\lim_{x\to 1} x\right]^4 + 7\left[\lim_{x\to 1} x\right] + 2\left[\lim_{x\to 1} 1\right] \\
&= 1 + 7 + 2 = 10.
\end{aligned}$$

Similarly, the denominator gives

$$\lim_{x\to 1} [x - 4] = 1 - 4 = -3.$$

Since both limits exist and the denominator is non-zero, we can apply the quotient Limit Law to get

$$\lim_{x\to 1} \frac{x^4+7x+2}{x-4} = \frac{\lim_{x\to 1} x^4+7x+2}{\lim_{x\to 1} x-4} = -\frac{10}{3}. \qquad \blacksquare$$

A similar argument to the previous example is the following theorem:

**Theorem 5.7**

If $f(x) = p(x)/q(x)$ is any rational functions (so that $p$ and $q$ are polynomials), and $c \in \mathbb{R}$ is such that $q(c) \neq 0$ then

$$\lim_{x\to c} \frac{p(x)}{q(x)} = \frac{p(c)}{q(c)}.$$

*Proof.* The key is to first show this for polynomials and apply the limit laws. We shall assume *a priori* that

$$\lim_{x\to c} x = c.$$

Now let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be an arbitrary polynomial. Certainly the limit of each $a_i x^i$ exists, and so the limit laws imply that

$$\begin{aligned}
\lim_{x\to c} p(x) &= \lim_{x\to c} \left[a_n x^n + \cdots + a_1 x + a_0\right] \\
&= a_n \left[\lim_{x\to c} x\right]^n + a_{n-1}\left[\lim_{x\to c} x\right]^{n-1} + \cdots + a_1 \left[\lim_{x\to c} x\right] + a_0 \\
&= a_n c^n + a_{n-1}c^{n-1} + \cdots + a_1 c + a_0 \\
&= p(c).
\end{aligned}$$

Thus the result holds for any polynomial. Now if $p$ and $q$ are two polynomials and $q(c) \neq 0$, then the limit laws for quotients implies

$$\lim_{x\to c} \frac{p(x)}{q(x)} = \frac{\lim_{x\to c} p(x)}{\lim_{x\to c} q(x)} = \frac{p(c)}{q(c)}. \qquad \square$$

**Example 5.8**

Determine the limit

$$\lim_{t\to 0} \frac{\sqrt{1+t}-\sqrt{1-t}}{t}.$$

*Solution.* When one finds that an expression involving square roots is proving difficult, it is often a good idea to multiply by a conjugate form which places the square roots into a more amenable position. In this instance, notice that

$$\left(\sqrt{1+t}-\sqrt{1-t}\right)\frac{\sqrt{1+t}+\sqrt{1-t}}{\sqrt{1+t}+\sqrt{1-t}} = \frac{(1+t)-(1-t)}{\sqrt{1+t}-\sqrt{1-t}} = \frac{2t}{\sqrt{1+t}+\sqrt{1-t}}.$$

Hence our limit becomes

$$\lim_{t\to 0} \frac{\sqrt{1+t}-\sqrt{1-t}}{t} = \lim_{t\to 0} \frac{2t}{t\left(\sqrt{1+t}+\sqrt{1-t}\right)} = \lim_{t\to 0} \frac{2}{\sqrt{1+t}+\sqrt{1-t}} = 1. \qquad \blacksquare$$

**Example 5.9**

Compute the limit

$$\lim_{h\to 0} \frac{(x+h)^2 - x^2}{h}$$

*Solution.* Notice here that the variable of the limit is $h$ rather than $x$! Since we cannot just substitute $h = 0$ into this equation, we must manipulate the expression to see if we can derive a more meaningful representative. Expanding the denominator we get

$$\begin{aligned} \lim_{h\to 0} \frac{(x+h)^2 - x^2}{h} &= \lim_{h\to 0} \frac{x^2 + 2xh + h^2 - x^2}{h} = \lim_{h\to 0} \frac{2xh + h^2}{h} \\ &= \lim_{h\to 0} 2x + h = 2x. \end{aligned} \qquad \blacksquare$$

## 5.4 Infinite Limits

There are two ways in which to consider "limits at infinity," either the function itself can diverge to infinity, or we can take a limit as $x \to \pm\infty$. The following sections discuss this behaviour.

### 5.4.1 Vertical Asymptotes

There are functions which are *singular* at a point $x = c$, and these can sometimes result in our limits being infinite. For example, the function $f(x) = 1/x$ becomes positively (resp. negatively) large when $x > 0$ (resp. $x < 0$) is small. So in this instance we have

$$\lim_{x\to 0^+} \frac{1}{x} = \infty \quad \text{and} \quad \lim_{x\to 0^-} \frac{1}{x} = -\infty. \tag{5.1}$$

Depending on the given function, the left and right limits might give the same signed infinity. For example,

$$\lim_{x\to 0} \frac{1}{x^2} = \infty.$$

In these instances we maintain that the limit *does not exist*, but will still write the equal sign to indicate that the one-sided limits agree.

**Definition 5.10**

The line $x = a$ is said to be a *vertical asymptote* for $f$ if either of the one sided limits is infinite; that is, if

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

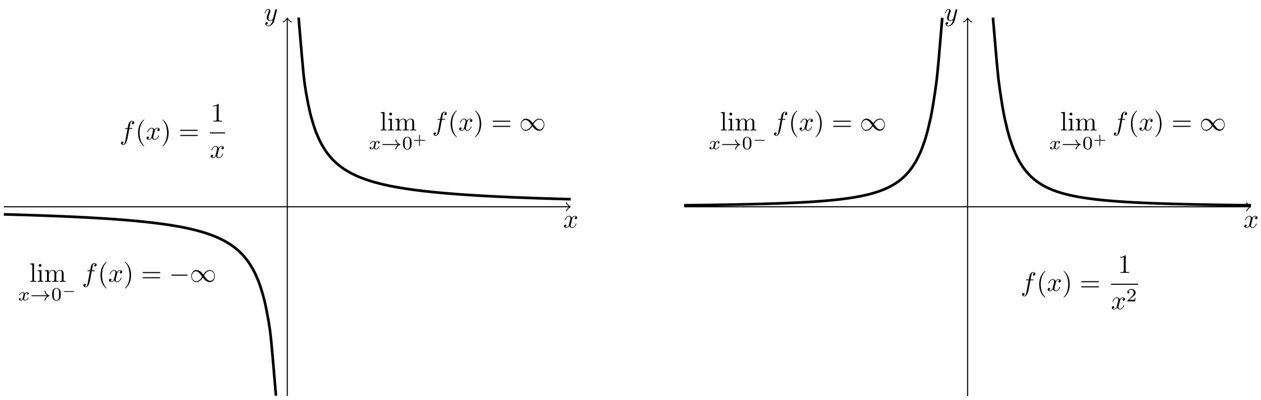



Figure 5.2: The functions $1/x$ and $1/x^2$ both have vertical asymptotes at $x = 0$.

**Example 5.11**

Determine the vertical asymptotes of the function $f(x) = \dfrac{1}{(x-3)(x-4)}$.

*Solution.* It will not take much to convince us that $f$ has vertical asymptotes at $x = 3$ and $x = 4$, since the numerator is constant but the denominator gets very small. The only question is whether or not this approaches positive or negative infinity. Notice that as $x \to 4$ we can always choose to limit ourselves to an arbitrarily small neighbourhood of $x = 4$, so in particular, let's assume that $3 < x < 4$. In this case $x - 3 > 0$ but $x - 4 < 0$, so that $(x-3)(x-4) < 0$. This tells us that

$$\lim_{x\to 4^-} \frac{1}{(x-3)(x-4)} = -\infty \quad \text{and} \quad \lim_{x\to 3^+} \frac{1}{(x-3)(x-4)} = -\infty.$$

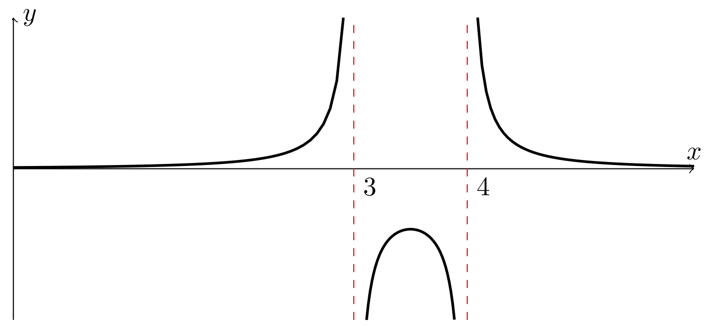
On the other hand, when $x > 4$ or $x < 3$, the same argument above implies that $(x-3)(x-4) > 0$, so that

$$\lim_{x\to 4^+} \frac{1}{(x-3)(x-4)} = \infty \quad \text{and} \quad \lim_{x\to 3^-} \frac{1}{(x-3)(x-4)} = \infty.$$

Thus the limit diverges to infinity in both instances, but to *different* infinities. ■

### 5.4.2 Horizontal Asymptotes

Many of the functions we have discussed so far fail to behave nicely as we tend to infinity. For example, the functions $x, x^2, e^x$ all become large as $x$ is allowed to grow large. However, there are some functions which exhibit a finite behaviour as we head off towards infinity, and we shall dedicate ourselves in the short term to examining such asymptotic behaviour.




$$f(x) = \frac{1}{(x-3)(x-4)}$$

Figure 5.3: The vertical asymptotes of the function found in Example 5.11.

**Definition 5.12**

If $f$ is defined on $(a, \infty)$ (resp. $(-\infty, a)$) then we say that

$$\lim_{x\to\infty} f(x) = L, \qquad \left(\text{resp. } \lim_{x\to-\infty} f(x) = L\right)$$

if whenever $x$ gets arbitrarily large and positive (resp. negative) then $f$ gets arbitrarily close to $L$. In such instances, we say that $L$ is a *horizontal asymptote* of $f$.

To re-iterate, polynomial functions are not going to have finite limits at infinity as they diverge off to infinity. So what are examples of functions which do have finite? Well to start, for any $p > 0$, we have

$$\lim_{x\to\pm\infty} \frac{1}{x^p} = 0.$$

This gives us a large number of unexciting functions with finite asymptotic behaviour, but serves as a stepping stone to deal with other, more exotic functions.

**Example 5.13**

Determine the limit $\displaystyle\lim_{x\to\infty} \frac{3x^2+6x-1}{4x^2-2}$.

*Solution.* It is unwieldy to deal with this function as written, since both the numerator and denominator become arbitrarily large and it is difficult to see what "cancellations" might occur. Instead, let's multiply and divide by the quantity $1/x^2$ to get

$$\begin{aligned}\lim_{x\to\infty} \frac{3x^2+6x-1}{4x^2-2}\frac{1/x^2}{1/x^2} &= \lim_{x\to\infty}\frac{3x^2/x^2+6x/x^2-1/x^2}{4x^2/x^2-2/x^2}\\ &= \lim_{x\to\infty}\frac{3+6/x-1/x^2}{4-2/x^2}\\ &= \frac{3+\lim\limits_{x\to\infty}\frac{6}{x}-\lim\limits_{x\to\infty}\frac{1}{x^2}}{4-\lim\limits_{x\to\infty}\frac{2}{x^2}}\\ &= \frac{3}{4}\end{aligned}$$

where in the last line we have used the fact that the quantities $\frac{1}{x}$ and $\frac{1}{x^2}$ go to zero as $x \to \infty$. ■

This suggests a general strategy for determining the limits of *rational functions*:

**Strategy for Rational Functions**

1. Determine the highest power $n$ which occurs in the functions,
2. Multiply and divide by $1/x^n$,
3. Take the limit as $x \to \infty$ using the fact that $1/x^p \to 0$ for all $p > 0$.

In the case of rational functions, this gives an incredibly convenient way of looking at a function and determining its asymptotic behaviour:

**Theorem 5.14**

Consider the rational function $f(x) = \dfrac{a_n x^n + \cdots + a_1 x + a_0}{b_m x^m + \cdots + b_1 x + b_0}$.

1. If $n > m$ then $\lim\limits_{x\to\infty} f(x) = \pm\infty$ (depending on the signs of $a_n$ and $b_m$).
2. If $n < m$ then $\lim\limits_{x\to\infty} f(x) = 0$
3. If $n = m$ then $\lim\limits_{x\to\infty} f(x) = \dfrac{a_n}{b_m}$.

As a matter of fact, the general strategy for determining the asymptotic nature of a function amounts to ascertaining which component of the function grows fastest. If the numerator grows fastest the function diverges, if the denominator grows fastest the function converges to 0, and if the numerator and denominator grow at the same rate, the function can attain a non-zero limit.

**Example 5.15**

Find the limit

$$\lim_{x\to\infty}\frac{e^x+e^{-x}}{e^x-e^{-x}}.$$

*Solution.* Our goal is to divide by the "highest order term." In this case, that corresponds to the $e^x$ term. Can you see why this is true? In essence, we would like to get rid of the things that explode as $x \to \infty$, and keep the things that get small. This means we want to get rid of $e^x$ and keep $e^{-x}$ which is why we divide by $e^x$. We thus get

$$\lim_{x\to\infty} \frac{e^x + e^{-x}}{e^x - e^{-x}} \frac{e^{-x}}{e^{-x}} = \lim_{x\to\infty} \frac{1 + e^{-2x}}{1 - e^{-2x}} = \frac{1}{1} = 1.$$ ■

**Exercise**

What are the limits

$$\lim_{x\to 0^\pm} \frac{e^x + e^{-x}}{e^x - e^{-x}}, \qquad \lim_{x\to -\infty} \frac{e^x + e^{-x}}{e^x - e^{-x}}.$$

**Example 5.16**

Determine the horizontal asymptotes of the function $f(x) = \dfrac{\sqrt{4x^2+5}}{x+2}$.

*Solution.* The square root here presents many difficulties. In particular, while the numerator will always be positive, the denominator will change sign. We must ensure that we account for this. Furthermore, while the numerator might have an $x^2$ component, the square root means that the numerator effectively acts as $\sqrt{x^2} = |x|$ so only grows linearly rather than quadratically.

In the limit as $x \to \infty$ we get

$$\begin{aligned}
\lim_{x\to\infty} \frac{\sqrt{4x^2+5}}{x+2} &= \lim_{x\to\infty} \frac{\sqrt{4x^2+5}}{x+2} \frac{1/x}{1/x} = \lim_{x\to\infty} \frac{\sqrt{4+5/x^2}}{1+2/x} \qquad \text{since } \frac{1}{x} = \frac{1}{\sqrt{x^2}} \text{ when } x > 0 \\
&= \frac{\sqrt{4 + \lim_{x\to\infty} \frac{5}{x^2}}}{1 + \lim_{x\to\infty} \frac{2}{x}} = \frac{\sqrt{4}}{1} = 2.
\end{aligned}$$

The tricky part above was that in order to pass the $1/x$ term into the square root, we needed to square it first. On the other hand, when we take the limit $x \to -\infty$, we will have that $\sqrt{x^2} = -x$ for $x < 0$, so that

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

Now a similar computation as that above yields

$$\lim_{x\to-\infty} \frac{\sqrt{4x^2+5}}{x+2} \frac{1/x}{1/x} = \lim_{x\to-\infty} -\frac{\sqrt{4+5/x^2}}{1+2/x} = -\frac{\sqrt{4}}{1} = -2.$$

Hence the horizontal asymptotes for $f$ occur at $\pm 2$. ■

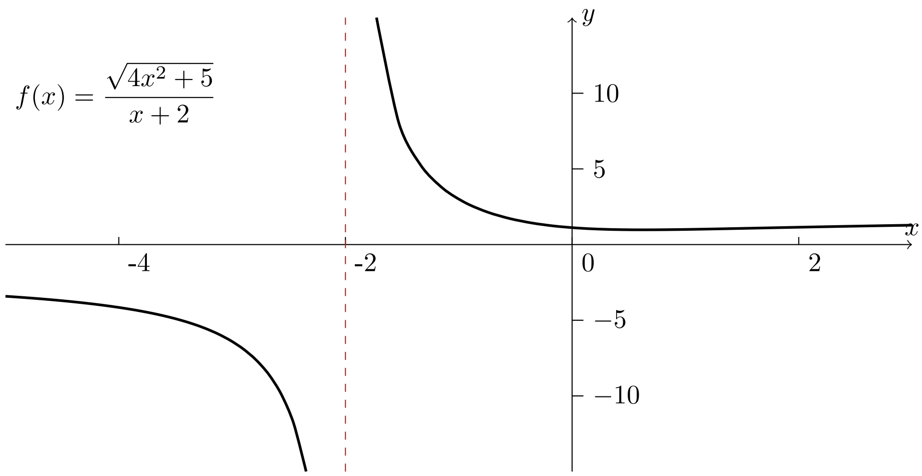



Figure 5.4: The plot of $f$ for Example 5.16

**Example 5.17**

Find the limit
$$\lim_{x\to\infty} \sqrt{x^4+10} - x^2.$$

*Solution.* At first glance, this may look like it diverges, but be careful! The fact that there is an $x^4$ term under a square root means $\sqrt{x^4+10}$ behaves as $x^2$ as $x$ becomes large. This term will, in the long term, cancel the effect of the $x^2$ leaving a finite answer. In order to make this more precise, we multiply by the conjugate

$$\lim_{x\to\infty}\left[\left(\sqrt{x^4+10}-x^2\right)\left(\frac{\sqrt{x^4+10}+x^2}{\sqrt{x^4+10}+x^2}\right)\right] = \lim_{x\to\infty}\frac{x^4+10-x^4}{\sqrt{x^4+10}+x^2} = \lim_{x\to\infty}\frac{10}{\sqrt{x^4+10}+x^2}.$$

The bottom term goes goes to $\infty$ as $x\to\infty$, so the whole limit goes to zero, and we conclude

$$\lim_{x\to\infty}\sqrt{x^4+10}-x^2 = 0. \qquad \blacksquare$$

## 5.5 Continuity

Of the examples seen thus far, there were instances in which the limit was computable by simply substituting the limiting value into the function, and other more pathological examples wherein this was not possible. The former examples are particularly special, not only because of the simplicity of evaluating limits, but because they tell us that the function is, in a sense, “well-behaved” at that limiting point.

**Definition 5.18**

A function $f$ is *continuous* at the point $c \in \mathbb{R}$ if

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

If the function $f$ is continuous at all points in in its domain, we say that it is *continuous*.

In Example 5.6 we showed that

$$\lim_{x \to 1} \frac{x^4 + 7x + 2}{x - 4} = -\frac{10}{3}.$$

Substituting $x = 1$ into this limiting function, we get

$$\frac{(1)^4 + 7(1) + 2}{(1) - 4} = \frac{10}{-3}.$$

so that $f(x) = (x^4 + 7x + 2)/(x - 4)$ is continuous at the point $x = 1$.

In fact, we have already seen entire families of continuous functions. Recall from Theorem 5.7 that if $p$ and $q$ are polynomials and $q(c) \neq 0$ then

$$\lim_{x \to c} \frac{p(x)}{q(x)} = \frac{p(c)}{q(c)}.$$

This implies that all rational functions are continuous at points where their denominator does not vanish.

In general, proving that a function is continuous is a complicated procedure which requires an in-depth knowledge of the properties of the function. Consequently, I'll omit the proofs that these functions are continuous. From here on, you are allowed to assume that the following functions are continuous on their domain:

1. All polynomials,
2. Root functions (example: $\sqrt{x}$, $\sqrt[3]{x}$),
3. Exponential functions $a^x$ for $a > 0$,
4. Logarithms $\log_a(x)$ for $a > 0$.

We can immediately deduce the following corollary from the Limit Laws

**Corollary 5.19**

If $f$ and $g$ are continuous at a point $c$, then

1. $\alpha f$ is continuous at $c$, for any real number $\alpha$,
2. $f \pm g$ is continuous at $c$,
3. $fg$ is continuous at $c$,
4. $f/g$ is continuous at $c$, provided $g(c) \neq 0$.

*Proof.* I will give the proof for the sum $f + g$. The remaining proofs all follow precisely the same argument, and are essentially derivative of the Limit Laws.

To show that $f + g$ is continuous, our goal is to show that

$$\lim_{x\to c}[f(x) + g(x)] = f(c) + g(c).$$

Since both $f(x)$ and $g(x)$ are continuous at $c$, we know by hypothesis that

$$\lim_{x\to c} f(x) = f(c) \text{and} \lim_{x\to c} g(x) = g(c).$$

In particular, both limits exist, so the Limit Laws imply that

$$\lim_{x\to c}[f(x) + g(x)] = \left[\lim_{x\to c} f(x)\right] + \left[\lim_{x\to c} g(x)\right] = f(c) + g(c)$$

which is precisely what we wanted to show. □

**Example 5.20**

Consider the function $f(x) = \dfrac{x^2 + 6x + 2}{(x+4)(x^2-1)}$. Determine the points where $f$ is not continuous.

*Solution.* This function fails to be continuous wherever its denominator vanishes. We may factor $(x + 4)(x^2 - 1) = (x + 4)(x - 1)(x + 1)$, which tells us that the points of discontinuity occur at $x = -4, -1, +1$. ■

Combining Corollary 5.19 with the list of functions assumed to be continuous, we can immediately generate very large families of continuous functions.

**Example 5.21**

Determine the limit $\displaystyle\lim_{x\to 4} \frac{x^2 + \sqrt{x}}{x^3 2^x}$.

*Solution.* The numerator $x^2 + \sqrt{x}$ is the sum of the continuous functions $x^2$ and $\sqrt{x}$, and hence is itself continuous. Similarly, the denominator $x^3 2^x$ is also product of continuous functions and hence is continuous. The quotient will be continuous so long as the denominator does not vanish at $x = 4$, and indeed

$$\lim_{x\to 4} x^3 2^x = 4^3 2^4 = 64 \times 16 \neq 0.$$

Thus the whole function $f(x) = (x^2 + \sqrt{x})/(x^3 2^x)$ is continuous at $x = 4$ and the limit can be evaluated by substitution:

$$\lim_{x\to 4} \frac{x^2 + \sqrt{x}}{x^3 2^x} = \frac{4^2 + \sqrt{4}}{(4)^3 2^4} = \frac{18}{64 \times 16} = \frac{9}{512}.$$ ■

**Composition of Functions:** Arguably, our most powerful operation on functions is that of composition, and it turns out that this will give us an insight into what it means to be continuous. Recall that $f$ is continuous at $c$ if the limit can be evaluated by simple substitution. This may be alternatively written as

$$\lim_{x\to c} f(x) = f(c) = f\left(\lim_{x\to c} x\right).$$

In a sense, it appears as though we are able to pass the limit inside of the function. In many ways, this is the true definition of continuity, so let's see how this might be useful to solving problems.

**Theorem 5.22**

If $\lim_{x\to c} g(x) = L$ and $f$ is continuous at $L$, then

$$\lim_{x\to c} f(g(x)) = f\left(\lim_{x\to c} g(x)\right) = f(L).$$

When $g(x) = x$ this reduces to the definition of continuity.

**Example 5.23**

Define the function

$$g(x) = \begin{cases} x & x \neq 0 \\ 2 & x = 0 \end{cases}.$$

Compute the limit $\lim_{x\to 0} e^{g(x)}$.

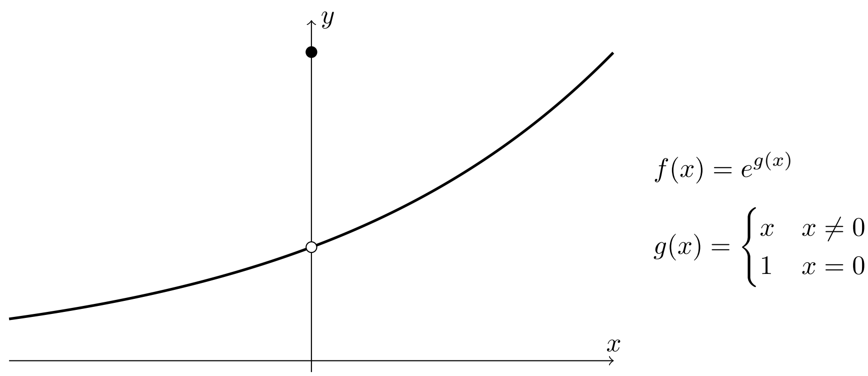



Figure 5.5: The function $f(x) = e^{g(x)}$, though $g$ is not a continuous function. Nonetheless, since $f$ *is* continuous, we can pass the limit inside the argument of $f$.

*Solution.* One might be tempted into thinking that $e^{g(x)}$ is continuous, but the fact that $g$ fails to be continuous means this is not the case. Indeed, if we attempted a direct substitution, we would find the limit would be $e^{g(0)} = e^2$, and this is not true (Figure 5.5). Instead, define $f(x) = e^x$ so that $e^{g(x)} = f(g(x))$. We know that $f$ is then continuous, so by the previous theorem we have

$$\lim_{x\to c} f(g(x)) = f\left(\lim_{x\to c} g(x)\right) = e^{\lim_{x\to c} g(x)} = e^0 = 1. \qquad \blacksquare$$

An immediate result of Theorem 5.22 is the following:

**Corollary 5.24**

If $g$ is continuous at $c$ and $f$ is continuous at $g(c)$ then $f \circ g$ is continuous at $c$.

*Proof.* Because both $f$ and $g$ are continuous at the necessary points, we just keep passing the limit into the arguments to see that

$$\lim_{x\to c} f(g(x)) = f\left(\lim_{x\to c} g(x)\right) = f\left(g\left(\lim_{x\to c} x\right)\right) = f(g(c)). \qquad \square$$

### 5.5.1 One-Sided Continuity and Failures of Continuity

Just as there are one sided limits, we can consequently have one-sided continuity.

**Definition 5.25**

We say that $f(x)$ is continuous at $c$ *from the right* (resp. from the left) if

$$\lim_{x\to c^+} f(x) = f(c) \qquad \left(\text{resp. } \lim_{x\to c^-} f(x) = f(c)\right).$$

Certainly, any function which is continuous at $c$ will be continuous from both the left and the right at $c$. For an example of a function which is only continuous from a single side, consider the function

$$f(x) = \begin{cases} 4x+2 & x \leq 0 \\ -x & x > 0 \end{cases}.$$

At $x = 0$, $f$ has the value $f(0) = 2$, and has one sided limits

$$\lim_{x\to 0^-} f(x) = 2 \quad \text{and} \quad \lim_{x\to 0^+} f(x) = 0.$$

This implies that $f$ is continuous from the left, as $\lim_{x\to 0^-} f(x) = f(0)$, but not from the right.

**Example 5.26**

Consider the function

$$f(x) = \begin{cases} x/|x| & \text{if } x \neq 0 \\ c & \text{if } x = 0 \end{cases}.$$

Determine the value of $c$ such that $f$ is continuous from the left at 0. What value of $c$ makes $f$ continuous from the right at 0?

*Solution.* All we need to do is determine the limit as $x \to 0^-$ and set $c$ to be this number. In this limit we may assume that $x < 0$, so that $|x| = -x$, and we get

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

so we set $c = -1$. Doing the same analysis for the limit as $x \to 0^+$ we get $c = 1$. ∎

If a function fails to have a limit at $c$, then it certainly has no chance of being continuous there either. On the other hand, the function $f(x) = (x-3)/(x^2 - x + 12)$ has a finite, two-sided limit at $x = 3$, but fails to be continuous there since $f$ is not defined at $x = 3$. The fact that some discontinuities seem inherently "worse" than other leads to a classification of discontinuities.

**Definition 5.27**

If $f$ is a function, define the one-sided limits

$$L_+ = \lim_{x \to c^+} f(x) \quad \text{and} \quad L_- = \lim_{x \to c^-} f(x).$$

If $f$ fails to be continuous at $c$, we say that $c$ is

1. A *removable discontinuity* if both $L_+$ and $L_-$ exist and $L_+ = L_-$.
2. A *jump discontinuity* if $L_+, L_-$ exist but $L_+ \neq L_-$.
3. An *essential discontinuity* if one of $L_\pm$ does not exist or is infinite.

**Example:**

1. The function $f(x) = (x-3)/(x^2 - x + 12)$ has a removable discontinuity at $x = 3$.
2. The function $f(x) = x/|x|$ has a jump discontinuity at $x = 0$.
3. The function $f(x) = 1/x$ has an essential discontinuity at $x = 0$.

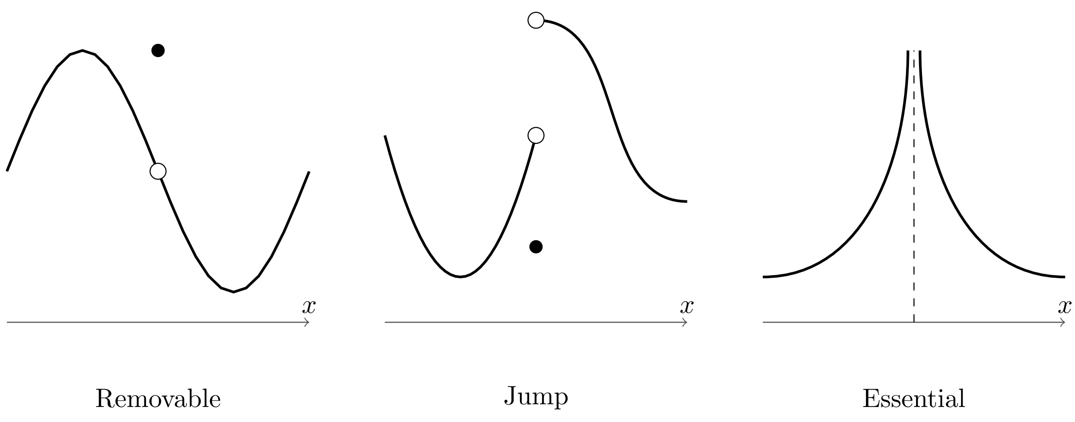



Figure 5.6: Examples of the types of discontinuity that can occur.

## 5.6 Exercises

5-1. Determine the following limits, if they exist:

(a) $\displaystyle\lim_{x\to 4} \frac{x^2-16}{x-4}$

(b) $\displaystyle\lim_{r\to 2} \frac{r^2+r-6}{r^2-4}$

(c) $\displaystyle\lim_{x\to -3} \frac{2-\sqrt{x^2-5}}{x+3}$

(d) $\displaystyle\lim_{x\to -1} \frac{\sqrt{x^2+8}-3}{x+1}$

(e) $\displaystyle\lim_{u\to 1} \frac{u^4-1}{u^3-1}$

(f) $\displaystyle\lim_{t\to 1} \frac{\sqrt{t}-t^2}{1-\sqrt{t}}$

5-2. Suppose that $\lim_{x\to -2} p(x) = 4$. $\lim_{x\to -2} r(x) = 0$, and $\lim_{x\to -2} s(x) = -3$. Find

(a) $\displaystyle\lim_{x\to -2} p(x)+r(x)+s(x)$

(b) $\displaystyle\lim_{x\to -2} p(x)r(x)s(x)$

(c) $\displaystyle\lim_{x\to -2} \frac{-4p(x)+5r(x)}{s(x)}$

(d) $\displaystyle\lim_{x\to -2} \frac{p(x)^2+s(x)^2+r(x)^2}{s(x)-p(x)}$

5-3. Find the limits

(a) $\displaystyle\lim_{x\to -0.5^-} \sqrt{\frac{x+2}{x+1}}$

(b) $\displaystyle\lim_{x\to -2^+} \left(\frac{x}{x+1}\right)\left(\frac{2x+5}{x^2+x}\right)$

(c) $\displaystyle\lim_{x\to 1^-} \left(\frac{1}{x+1}\right)\left(\frac{x+6}{x}\right)\left(\frac{3-x}{7}\right)$

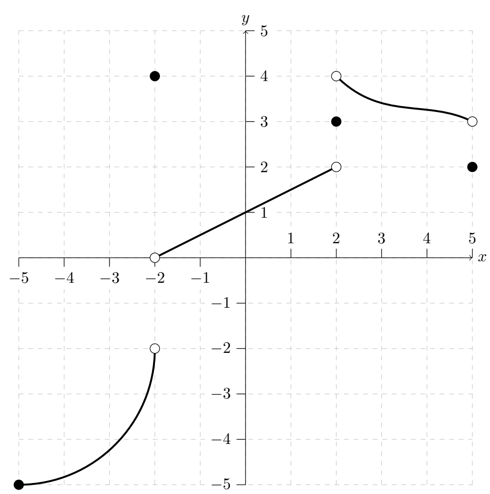
5-4. Determine the following one-sided limits:

(a) $\displaystyle\lim_{t\to 3^\pm} \frac{t^2+4t-21}{t-3}$

(b) $\displaystyle\lim_{x\to 0^\pm} \frac{|x|}{x}$

(c) $\displaystyle\lim_{x\to 1^\pm} f(x)$ where $f(x) = \begin{cases} x^2+2x & x<1 \\ 3x+4 & x\geq 1 \end{cases}$.

5-5. Consider the following graph:



Determine the following:

(a) $\lim_{x\to 0} f(x)$

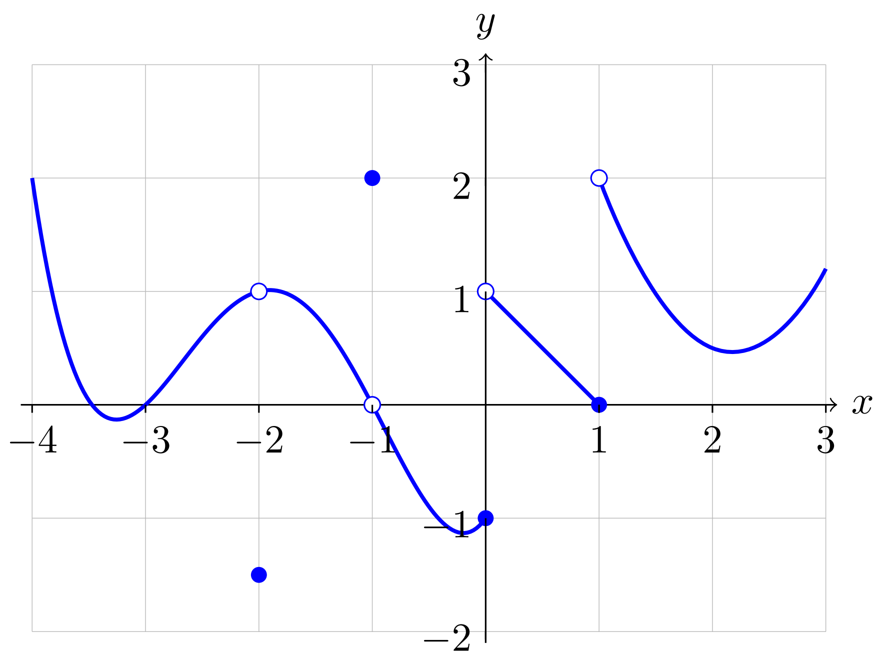
(b) $\lim_{x\to -2^+} f(x)$

(c) $\lim_{x\to -2^-} f(x)$

(d) $\lim_{x\to 2} f(x)$

(e) $\lim_{x\to 0} f(|x|+2)$

5-6. Consider a function $f$ whose graph on $[-4, 3]$ is given below:




Determine each limit, if it exists:

(a) $\lim_{x\to 0^-} f(x)$

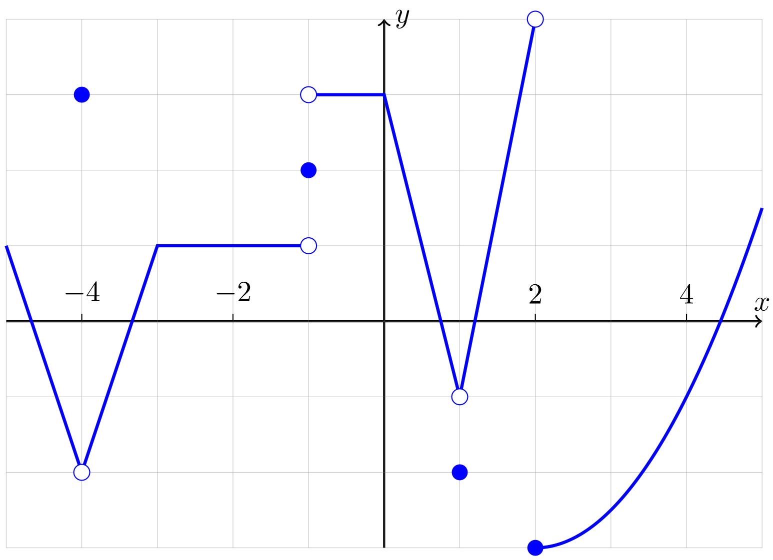
(b) $\lim_{x\to -1} f(x)$

(c) $\lim_{x\to 1} f(x)$

(d) $\lim_{x\to -3^+} f(f(x))$

(e) $\lim_{x\to 0} f(|x|+1)$

5-7. The graph of the function $f$ on the interval $[-5, 5]$ is given below. Determine each of the given limits.




(a) $\lim_{x\to -4} f(x)$

(b) $\lim_{x\to -1^+} f(x)$

(c) $\lim_{x\to 0} f(\sqrt{4+x^2})$

(d) $\lim_{x\to 1} f(-e^{x-1})$

(e) $\lim_{x\to -4} f(f(x))$

(f) $\lim_{x\to -2} f(f(x))$

(g) $\lim_{x\to 1} f(f(x))$

(h) $\lim_{x\to -3} f(f(x))$

5-8. True or False:

(a) If $\lim_{x\to c} f(x) = L$, then $f(c) = L$.

(b) If $\lim_{x\to c}[f(x)+g(x)]$ exists, then both $\lim_{x\to c} f(x)$ and $\lim_{x\to c} g(x)$ exist.

(c) If both $\lim_{x\to c} f(x)$ and $\lim_{x\to c} g(x)$ exist, then $\lim_{x\to c}[f(x)+g(x)]$ exists

(d) If $\lim_{x\to c} f(x) = L$, then $\lim_{x\to c} |f(x)| = L$.

5-9. Determine the vertical asymptotes of the following functions. In each case, determine whether the corresponding one-sided limits diverge to $\infty$ or $-\infty$.

(a) $f(x) = \dfrac{1}{x^2+5x+6}$

(b) $f(x) = \dfrac{1}{x} + \dfrac{1}{x-1} + \dfrac{1}{x-2}$

(c) $f(x) = \log(x)$

(d) $f(x) = \dfrac{1}{\log(x)-1}$

(e) $f(x) = e^{1/x}$

5-10. Determine the horizontal asymptotes of the following functions:

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

(b) $f(x) = \dfrac{-2x^2+3x+15}{4x^2-x+6}$

(c) $f(x) = \dfrac{\sqrt{2x^2+1}}{x+6}$

(d) $f(x) = \sqrt{9x^2+1} - 3x$

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

5-11. Compute each of the following limits. Where the limit is $\infty$, indicate whether it is $+\infty$ or $-\infty$.

(a) $\displaystyle\lim_{x\to 0^+} \frac{1}{|x|}$

(b) $\displaystyle\lim_{s\to 1^-} \frac{1}{\ln(s)}$

(c) $\displaystyle\lim_{a\to 0^+} \frac{2}{e^a - e^{-a}}$

(d) $\displaystyle\lim_{k\to 0^+} \frac{k^2-2k-15}{k^2-k-12}$

(e) $\displaystyle\lim_{z\to\infty} \frac{e^z+e^{-z}}{2}$

(f) $\displaystyle\lim_{t\to\infty} \left[t+\sqrt{t^2-1}\right]$

(g) $\displaystyle\lim_{m\to-\infty} \left[m+\sqrt{m^2-1}\right]$

(h) $\displaystyle\lim_{u\to\infty} \left[u-\sqrt{u^2-1}\right]$

(i) $\displaystyle\lim_{r\to-\infty} \frac{e^r+e^{-r}}{e^r-e^{-r}}$

(j) $\displaystyle\lim_{v\to\infty} \frac{e^v+e^{2v}}{e^{3v}+e^{4v}}$

5-12. Find the values of $a, b$ which make $f$ everywhere continuous:

(a) $f(x) = \begin{cases} x^2+1 & x < -1 \\ ax+b & -1 \leq x \leq 1 \\ |x-2| & x > 1 \end{cases}$

(b) $f(x) = \begin{cases} ax+2b & x \leq 0 \\ x^2+3a-b & 0 < x \leq 2 \\ 3x-5 & x > 2 \end{cases}$

5-13. For each function, determine where the function fails to be continuous. Classify the discontinuity.

(a) $f(x) = \dfrac{x^3-3x}{x-3}$

(b) $f(x) = \dfrac{x}{|x|}$

(c) $f(x) = \dfrac{x^3-3x}{x-3}$

(d) $f(x) = \dfrac{3x-4}{x^2-2x-15}$

(e) $f(x) = x\log(x)$

(f) $f(x) = \dfrac{1}{\sqrt{1-x^2}}$

(g) $f(x) = \begin{cases} (x-3)^2+5 & x < 2 \\ 2x+3 & x \geq 2 \end{cases}$

5-14. True or False:

(a) If $f$ is everywhere continuous, then $|f|$ is everywhere continuous.

(b) If $|f|$ is everywhere continuous, then $f$ is everywhere continuous.

(c) If $f+g$ is continuous at $c$, then both $f$ and $g$ are continuous at $c$.

(d) If both $f$ and $g$ are continuous at $c$, then $f+g$ is continuous at $c$.

(e) If $g$ is continuous at $c$, and $f$ is continuous at $g(c)$, then $\lim_{x\to c} f(g(x)) = f(g(c))$.

(f) If $\lim_{x\to 3} f(x) = 4$ and $f(3) = 3$ then $f$ has a removable discontinuity at $x = 3$?

(g) If $f$ is not continuous at $c$, then $f(c)$ is not defined.

5-15. In each part below, a series of properties are given. You must give an *algebraic* description of a function which satisfies *all* of the properties simultaneously.

(a) A function $f$ which is continuous on all of $\mathbb{R}$.

(b) A function $f$ which is continuous on all of $\mathbb{R}$ except the points $x = 5$, where it has a jump discontinuity, and $x = 7$ where it has an essential discontinuity.

(c) A function $f$ which is discontinuous at exactly three different points, has a horizontal asymptote as $x \to \infty$, but does not have a horizontal asymptote as $x \to -\infty$.

(d) A function $f$ with two horizontal asymptotes, and two vertical asymptotes.

(e) A function $f$ which is discontinuous at every integer.

5-16. You work for a toy company, preparing for production of a new product called Lincoln Logarithms. You are manufacturing the toys in-house, with each toy requiring \$4 in material, in addition to a one-time fee of \$25,000 to establish the assembly line. As your workers produce more toys, they become more efficient, making your labour costs $L(q) = 30000/q$ dollars, where $q$ is the number of toys you produce.

(a) Let $T(q)$ be the total cost and $A(q)$ be the average cost to produce $q$ toys. Write down formulas for $T$ and $A$.

(b) Determine the limit $\lim_{q\to\infty} A(q)$.

5-17. In economics, *utility* is a measure of satisfaction or usefulness. The utility of wealth could therefore be thought of as how happy you are as a function of the amount of wealth you possess. When applied to economic measurements, the word "*marginal*" describes the change corresponding to an increase in the measured quantity. The marginal utility of wealth is how much your happiness increases as your wealth increases. The *Law of Diminishing Marginal Utility* states that as you consume more, your marginal utility decreases; that is, the more you have of something, the smaller the impact getting more of it affects your happiness. If $w$ is your wealth, and $T(w)$ is your utility of wealth, your marginal utility of wealth at $w$ is

$$M(w) = \lim_{h\to 0} \frac{T(w+h) - T(w)}{h},$$

and the rate at which your marginal utility of wealth is changing at $w$ is

$$C(w) = \lim_{h\to 0} \frac{M(w+h) - M(w)}{h}.$$

Suppose $w$ is measured in 10s of thousands of dollars, and $T(w) = \dfrac{300w}{w+1}$ is the utility of your wealth.

(a) Find a closed form expression for your marginal utility of happiness; that is, find a formula for $M(w)$ which does not contain a limit.

(b) Find a closed form for $C(w)$ and confirm that this number is always negative when $w > 0$, hence showing that the marginal utility of wealth is diminishing.

# 6 Derivatives

The power of calculus is that it gives us to tools to analyze how things change in time, or more precisely the instantaneous rate of change. This will be done by exploiting properties of limits.

## 6.1 First Principles

In Section 5 we considered the problem of measuring the speed of a race car. This was done by measuring the distance travelled by the car over successively smaller time intervals. While this was our motivation for introducing limits, we never finished the example and actually applied a limit to the problem. We now remedy that oversight.

Let $f(t)$ represent the position of the car at time $t$, and let $t_f$ denote the time at which the car passes the finish line. The average speed of the car in the second after passing the finish line is given by

$$\begin{matrix}\text{average speed} \\ \text{after 1 second}\end{matrix} = \frac{\text{distance travelled}}{\text{time elapsed}} = \frac{f(t_0 + 1\text{ sec}) - f(t_0)}{1\text{ sec}}.$$

More generally by letting $h$ denote an arbitrary quantity of time the average speed,

$$\begin{matrix}\text{average speed} \\ \text{after } h \text{ seconds}\end{matrix} = \frac{\text{distance}}{\text{time}} = \frac{f(t_0 + h) - f(t_0)}{h},$$

where if $h > 0$ then we are measuring after crossing the finish line, and if $h < 0$ we are measuring prior to crossing the finish line. The instantaneous speed of the car may then be determined by taking the limit as our time interval becomes arbitrarily small; that is, by taking $h \to 0$:

$$\begin{matrix}\text{instantaneous speed} \\ \text{at time } t_0\end{matrix} = \lim_{h\to 0} \frac{f(t_0 + h) - f(t_0)}{h}.$$

There is no reason why we cannot generalize this discussion to a more abstract setting or to other examples. The key point is that having precise knowledge about a function $f$ near a given point $t_0$ allows us to determine how quickly $f$ is changing at $t_0$.

**Definition 6.1**

Let $f$ be a function and $x_0$ a point in the domain of $f$. If the limit

$$f'(x_0) = \lim_{h\to 0} \frac{f(x_0 + h) - f(x_0)}{h} \tag{6.1}$$

exists, we say that $f$ is *differentiable at* $x_0$ and $f'(x_0)$ is the *derivative of* $f$ *at* $x_0$. If (6.1) exists for all $x$ in the domain of $f$, we simply say that $f$ is *differentiable*.

**Example 6.2**

Compute the derivative of $f(x) = x$ and $g(x) = x^2$ at the points $x = -1$ and $x = 5$.

*Solution.* With nothing more than Definition 6.1 to work with, we set to work. For the linear function $f(x) = x$, at $x = -1$ we get

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

while at $x = 5$ we find

$$f'(5) = \lim_{h\to 0} \frac{f(5+h) - f(5)}{h} = \lim_{h\to 0} \frac{(5+h) - 5}{h} = 1.$$

This shows that the derivative of $f(x) = x$ is the same at both $x = -1$ and $x = 5$. In fact, it is not too hard to convince ourselves that regardless of what value we substitute for $x_0$, we are always going to get $f'(x_0) = 1$. This is tantamount to the fact that $f'(x_0)$ describes the rate of change of the function, and $f(x) = x$ is function which is growing at a constant rate.

Before proceeding to the next example, let us try to anticipate the kind of solution we will see from the function $g(x) = x^2$. The shape of the parabola seems to suggest that $x^2$ is decreasing at $x = -1$ and so $g'(-1)$ should be negative. On the other hand, $g(x)$ is growing larger at $x = 5$ implying that $g'(5)$ should be positive. The necessary computations are slightly more complicated, but nothing beyond our capabilities:

$$\begin{aligned}
g'(-1) &= \lim_{h\to 0} \frac{g(-1+h) - g(-1)}{h} = \lim_{h\to 0} \frac{(-1+h)^2 - (-1)^2}{h} \\
&= \lim_{h\to 0} \frac{(1 - 2h + h^2) - 1}{h} = \lim_{h\to 0} -2 + h \\
&= -2 \\
g'(5) &= \lim_{h\to 0} \frac{g(5+h) - g(5)}{h} = \lim_{h\to 0} \frac{(5+h)^2 - 5^2}{h} \\
&= \lim_{h\to 0} \frac{h^2 + 10h + 25 - 25}{h} = \lim_{h\to 0} h + 10 \\
&= 10.
\end{aligned}$$

This is precisely the behaviour we anticipated: The rate at which $g$ is growing at $x = 5$ is much faster than the speed at which it is decreasing at $x = -1$, and is corroborated by the graph of $g(x) = x^2$. ■

**Example 6.3**

Compute the derivative of the function $f(x) = \dfrac{3x+2}{x-4}$ at the point $x = 5$.

*Solution.* The behaviour of this function is not as transparent as that given in Example 6.2, so we must trust in our calculations. As before, we apply Definition 6.1 to $f$ to find

$$\begin{aligned}
f'(5) &= \lim_{h\to 0} \frac{f(5+h) - f(5)}{h} = \lim_{h\to 0} \frac{\frac{3(5+h)+2}{(5+h)-4} - 17}{h} \\
&= \lim_{h\to 0} \frac{\frac{3h+17}{h+1} - 17}{h} = \lim_{h\to 0} \frac{\frac{(3h+17)-(17h+17)}{h+1}}{h} && \text{common denominator} \\
&= \lim_{h\to 0} \frac{-14h}{h(h+1)} = \lim_{h\to 0} \frac{-14}{h+1} \\
&= -14.
\end{aligned}$$

This tells us that the function is decreasing at $x = 5$. ■

The computations of the derivatives in Example 6.2 were redundant, requiring only small changes between the $x = -1$ and $x = 5$ cases. If instead of using a particular number, we let $x = a$ describe a general point, we could compute

$$\begin{aligned} f'(a) &= \lim_{x\to a} \frac{f(x)-f(a)}{x-a} = \lim_{x\to a}\frac{x^2-a^2}{x-a} \\ &= \lim_{x\to a}\frac{(x+a)(x-a)}{x-a} = \lim_{x\to a}(x+a) = 2a. \end{aligned}$$

This agrees with what we found when $a = -1$ and $a = 5$, but gives us the derivative at any point $a$. Hence one can view the derivative of a function as a function itself, with

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

In fact, this is where the word *derivative* comes from – the value of the function $f'(x)$ is derived from that of $f(x)$.

**Exercise:** How is the domain of $f'$ related to $f$? Must they always be the same? Or can one be larger than the other?

### 6.1.1 The Geometry of the Derivative

We can visualize the derivative geometrically by examining the graph of $f$.

**Definition 6.4**

Consider the graph of a function $f$ and let $a < b$ be distinct real numbers. The *secant line from $a$ to $b$* is the unique straight line which passes through the points $(a, f(a))$ and $(b, f(b))$.

Of interest is the slope of this secant line, given by

$$m_{ab} = \frac{f(b)-f(a)}{b-a}.$$

In fact, this slope describes the average change in the value of $f$ between $f(a)$ and $f(b)$. We can get successively better approximations to the instantaneous rate of change of a function by taking the distance between $a$ and $b$ to be successively smaller, and the instantaneous rate of change will be given by taking a limit as $a \to b$. See Figure 6.1. In this limit the secant line becomes a *tangent line*; that is,

If $f$ is differentiable at $a$, the slope of the tangent line to the graph of $f$ at $a$ is $f'(a)$.

Since we have a slope and a point in the plane $(a, f(a))$, we can form the equation of the tangent line through $f$ at $a$ using the point-slope formula:

$$y - f(a) = f'(a)(x-a). \tag{6.2}$$

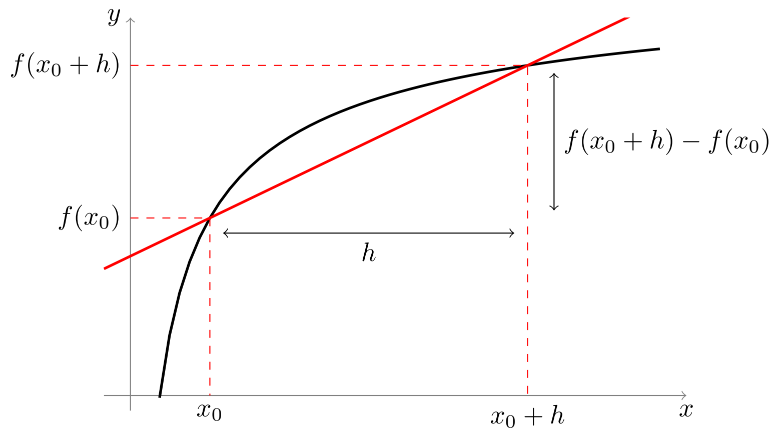



Figure 6.1: The difference quotient $[f(x_0+h)-f(x_0)]/h$ is the slope of the secant line between $(x_0, f(x_0))$ and $(x_0+h, f(x_0+h))$. In the limit as $h \to 0$, this becomes the slope of the tangent line.

**Example 6.5**

Determine the equation of the tangent line through $f(x) = \sqrt{x}$ at the point $x = 1$.

*Solution.* To determine the slope we need the value $f'(1)$, though it is not extra work to compute the derivative in general.

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = \lim_{h\to 0} \frac{\sqrt{x+h}-\sqrt{x}}{h} \\ &= \lim_{h\to 0} \frac{\sqrt{x+h}-\sqrt{x}}{h}\left[\frac{\sqrt{x+h}+\sqrt{x}}{\sqrt{x+h}+\sqrt{x}}\right] && \text{Multiply by conjugate} \\ &= \lim_{h\to 0} \frac{x+h-x}{h(\sqrt{x+h}+\sqrt{x})} \\ &= \lim_{h\to 0} \frac{1}{\sqrt{x+h}+\sqrt{x}} = \frac{1}{2\sqrt{x}}. \end{aligned}$$

At $x = 1$ we have $f(1) = \sqrt{1} = 1$, so we know our line passes through the point $(1, 1)$. Our derivative formula tells us that $f'(1) = \frac{1}{2}$, so using our point slope formula, we thus deduce that the equation of the tangent line is $y - 1 = (x - 1)/2$ or $y = x/2 + 1/2$. ■

**Example 6.6**

Find the equation of the tangent line to the graph $f(x) = 1/x$ at the point $x = 1$.

*Solution.* While (6.2) tells us that it is sufficient to find $f'(1)$, it is not much extra work to find

$f'(x)$ in general:

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = \lim_{h\to 0} \frac{\frac{1}{x+h}-\frac{1}{x}}{h} \\ &= \lim_{h\to 0} \frac{x-(x+h)}{hx(x+h)} = \lim_{h\to 0} \frac{-1}{x(x+h)} \\ &= -\frac{1}{x^2}, \end{aligned}$$

hence $f'(1) = -1$. To find the equation of the line, we need only determine a point through which the line passes. Since the line is tangent to $f$ at $x = 1$, it must pass through the point $(1, f(1)) = (1, 1)$. In conclusion, (6.2) tells us that the tangent line is

$$y - 1 = -1(x-1) \quad \text{or equivalently} \quad y = -x + 2. \qquad \blacksquare$$

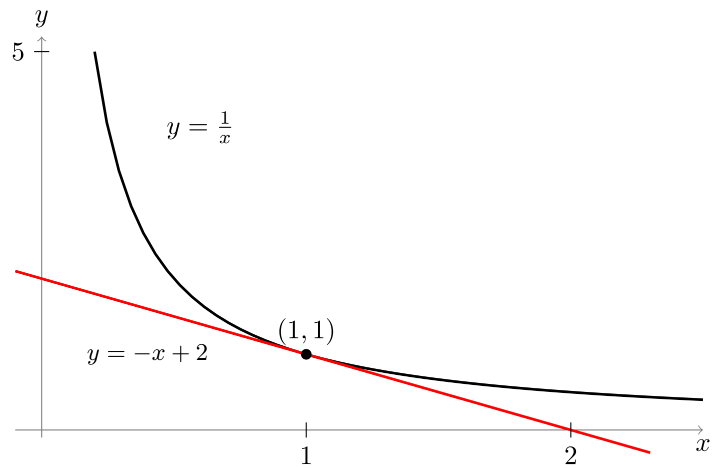



Figure 6.2: The derivative of $f(x) = 1/x$ at the point $x = 1$ gives us the slope of the tangent line to the graph at $x = 1$. Combining this with the fact that the line must pass through the point $(1, f(1)) = (1, 1)$ we know precisely the equation of the tangent line.

### 6.1.2 A Different Parameterization

An alternative way of writing the derivative comes from changing how we parameterize the limit. Recall from Definition 6.1 that

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

where the parameter $h$ denotes the distance from $a$. Taking the limit $h \to 0$ corresponds to letting our point $a + h$ get close to $a$. If we instead set $x = a + h$, the condition $h \to 0$ is equivalent to $x \to a$, giving an equivalent definition of the derivative:

$$f'(a) = \lim_{x\to a} \frac{f(x)-f(a)}{x-a} \tag{6.3}$$

**Example 6.7**

Compute the derivative of the function $f(x) = 9/x$ for any point $a \neq 0$ using (6.3).

*Solution.* Applying Equation (6.3) we get

$$\begin{aligned} f'(a) &= \lim_{x\to a} \frac{f(x) - f(a)}{x - a} = \lim_{x\to a} \frac{9/x - 9/a}{x - a} \\ &= \lim_{x\to a} \frac{9(a-x)/xa}{x-a} = \lim_{x\to a} \frac{-9}{ax} \\ &= -\frac{9}{a^2}. \end{aligned}$$

■

### 6.1.3 Relating Variables and Leibniz Notation

Functions can be used to describe how one variable depends upon another. For example, the identity $y = x^2$ tells us precisely how the variable $y$ depends upon the function $x$. Given a general function $f$, a functional relation between $y$ and $x$ can be written as $y = f(x)$. In this context, the derivative $f'(x)$ describes the instantaneous rate of change of $y$ as a function of $x$.

When the relationship is defined explicitly, say $y = x^2$, we need a way of talking about the derivatives without referring to the defining function $f(x) = x^2$. To do this, we write

$$\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{x=c} = f'(c) \qquad \text{or just} \qquad \frac{\mathrm{d}y}{\mathrm{d}x} \qquad \text{when we are feeling lazy.}$$

The notation $\frac{\mathrm{d}y}{\mathrm{d}x}$ represents how the variable $y$ is changing with respect to $x$. The motivation for this notation comes from the definition of the derivative as the limit of secant lines. If $a < b$ are two distinct real numbers, the change in the $y$-value of the function $f$ between $a$ and $b$ may be written as $\Delta y = f(b) - f(a)$ (this is still often used amongst physicists), with the change in the $x$-value being written as $\Delta x = b - a$. The secant line between $a$ and $b$ is then

$$m_{ab}^f = \frac{\Delta y}{\Delta x} \qquad \text{“} \xrightarrow{\Delta x \to 0} \text{”} \qquad \frac{\mathrm{d}y}{\mathrm{d}x},$$

and in the limit as $\Delta x \to 0$ these delta's “transform” into $d$'s.

**Example 6.8**

Determine $\dfrac{\mathrm{d}y}{\mathrm{d}x}$ if $y = (x^2 + 1)/(2x)$.

*Solution.* Applying the definition of the derivative, one finds that

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= \lim_{h\to 0} \frac{\frac{(x+h)^2+1}{2(x+h)} - \frac{x^2+1}{2x}}{h} = \lim_{h\to 0} \frac{\frac{(x^2+2xh+h^2)(2x)-(x^2+1)(2x+2h)}{(2x)(2x+2h)}}{h} \\
&= \lim_{h\to 0} \frac{(2x^3+4x^2h+2x)-(2x^3-2x^2h-2x-2h)}{h(2x)(2x+2h)} \\
&= \lim_{h\to 0} \frac{2x^2h-2h}{h(2x)(2x+h)} = \frac{2x^2-2}{4x^2} = \frac{x^2-1}{2x^2}.
\end{aligned}$$

■

Leibniz notation lets us think of differentiation as an *operator*; that is, something which acts on functions to create new functions. If $f$ is a function, we will use the notation

$$\frac{\mathrm{d}}{\mathrm{d}x} f(x)$$

to represent the action of taking the derivative of $f$ with respect to $x$.

## 6.2 Some Derivative Results

Similar to how the limit laws simplified the process of taking limits, we wish to establish a collection of tools to simplify the process of differentiation. In this section, we give a few formulas to simplify some calculations, and examine how to differentiate polynomials and exponential functions.

### 6.2.1 Linearity and the Power Rule

Taking derivatives plays nicely with addition and multiplication by a constant, as the following proposition illustrates:

**Proposition 6.9**

If $f$ and $g$ are differentiable at $c$, then

1. For any constant $\alpha \in \mathbb{R}$ the function $(\alpha f)(x) = \alpha f(x)$ is differentiable at $c$, and moreover
$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\alpha f(x)\right] = \alpha f'(x).$$
2. The function $(f+g)(x) = f(x) + g(x)$ is differentiable at $c$, and moreover
$$\frac{\mathrm{d}}{\mathrm{d}x}\left[f(x)+g(x)\right] = f'(x) + g'(x).$$

*Proof.* 1. By the definition of the derivative, we have

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\alpha f(x)\right] = \lim_{h\to 0} \frac{\alpha f(x+h) - \alpha f(x)}{h} = \alpha \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = \alpha f'(x).$$

2. Again, using the definition of the derivative we have

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\left[f(x)+g(x)\right] &= \lim_{h\to 0}\frac{[f(x+h)+g(x+h)]-[f(x)+g(x)]}{h} \\
&= \lim_{h\to 0}\frac{[f(x+h)-f(x)]+[g(x+h)-g(x)]}{h} && \text{re-arranging terms} \\
&= \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}+\lim_{h\to 0}\frac{g(x+h)-g(x)}{h} && \text{by the limit laws} \\
&= f'(x)+g'(x). && \square
\end{aligned}
$$

The previous theorem tells us that $\frac{\mathrm{d}}{\mathrm{d}x}$ is what is called a *linear operator*, in that it preserves scalar multiplication and addition. As an immediate corollary, we see that

$$
\frac{\mathrm{d}}{\mathrm{d}x}\left[f(x)-g(x)\right] = \frac{\mathrm{d}}{\mathrm{d}x}\left[f(x)+(-1)\times g(x)\right] = f'(x)+(-1)\times g'(x) = f'(x)-g'(x)
$$

so that the derivative of the difference is also the difference of the derivatives.

**Proposition 6.10**

For any positive integer $n$, $\dfrac{\mathrm{d}}{\mathrm{d}x}x^n = nx^{n-1}$.

*Proof.* Recall that one may always factor the $n^{th}$-powered difference of two elements as

$$
(a^n-b^n) = (a-b)(a^{n-1}+a^{n-2}b^1+\cdots+a^1b^{n-2}+b^{n-1}).
$$

Applying this to the definition of the derivative, we see that

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}x^n &= \lim_{y\to x}\frac{y^n-x^n}{y-x} = \lim_{y\to x}\frac{(y-x)(y^{n-1}+y^{n-2}x+\cdots+yx^{n-2}+x^{n-1})}{y-x} \\
&\lim_{y\to x}\underbrace{y^{n-1}+y^{n-2}x+\cdots+yx^{n-2}+x^{n-1}}_{n\text{-times}} \\
&= nx^{n-1}. && \square
\end{aligned}
$$

Combining Propositions 6.9 and 6.10 allow us to differentiate polynomials very quickly. Every polynomial function is built from scalar multiplication and addition of monomials of the form $a_nx^n$, so

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}x}\left[a_nx^n+a_{n-1}x^{n-1}+\cdots+a_2x^2+a_1x+a+0\right] \\
&= a_n\left[\frac{\mathrm{d}}{\mathrm{d}x}x^n\right]+a_{n-1}\left[\frac{\mathrm{d}}{\mathrm{d}x}x^{n-1}\right]+\cdots+a_1\left[\frac{\mathrm{d}}{\mathrm{d}x}x\right]+0 \\
&= na_nx^{n-1}+(n-1)a_{n-1}x^{n-2}+\cdots+2a_2x+a_1.
\end{aligned}
$$

**Example 6.11**

Compute the derivative of $x^{654}+13x^{45}-20x^5+6$.

*Solution.* Using our template above, we see that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\left[x^{654}+13x^{45}-20x^5+6\right] &= (654)x^{653}+13(45)x^{44}-20(5)x^4 \\ &= 654x^{653}+585x^{44}-100x^4. \qquad \blacksquare\end{aligned}$$

While the above proposition only showed that $\frac{\mathrm{d}}{\mathrm{d}x}x^n = nx^{n-1}$ for any *positive integer* $n$, it turns out that the formula works for all real numbers. We will see why this is the case in Section 11.3.1.

### 6.2.2 The Natural Exponent

One way of defining Euler's number $e$, defined in Section 1.5.3, is as the unique number which satisfies

$$1 = \lim_{h\to 0}\frac{e^h-1}{h}; \tag{6.4}$$

that is, it is the unique number $a$ for which the slope of the tangent line to $f(x) = a^x$ at $x = 0$ is precisely 1. Using this, the derivative of $e^x$ is computed as follows:

**Proposition 6.12**

The exponential function $e^x$ is its own derivative; that is, $\dfrac{\mathrm{d}}{\mathrm{d}x}e^x = e^x$

*Proof.* By definition of the derivative,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}e^x &= \lim_{h\to 0}\frac{e^{x+h}-e^x}{h} = \lim_{h\to 0}\frac{e^xe^h-e^x}{h} \\ &= e^x\lim_{h\to 0}\frac{e^h-1}{h} && \text{by (6.4)} \\ &= e^x. \qquad \square\end{aligned}$$

A similar argument will show that for any $a > 0$, $\frac{\mathrm{d}}{\mathrm{d}x}a^x = Ka^x$ for some $K$, which satisfies $K = \left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_{x=0} a^x$. We will later see that $K = \ln(a)$, but for the moment we do not have the tools to make this clear.

**Exercise:** Recall that if $n$ is a positive integer, we define $n! = n(n-1)(n-2)\cdots(3)(2)$. For example, $3! = 3\times 2 = 6$ and $4! = 4\times 3\times 2 = 24$. Consider the function

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

Compute the derivative of $f$. If you are given that derivatives are unique up to additive constants, what function must $f$ be?

### 6.2.3 The Product and Quotient Rule

Computing the derivative of sums of functions was ultimately rather simple. However, it turns out that computing the derivative of a product is a far more complicated affair.

**Theorem 6.13**

If $f$ and $g$ are differentiable at $c$, then $fg$ is differentiable at $c$ and

$$\left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_{x=c} f(x)g(x) = f'(c)g(c) + g'(c)f(c).$$

*Proof.* By the limit definition of the derivative, we have

$$\begin{aligned}
\left.\frac{\mathrm{d}}{\mathrm{d}c}\right|_{c=c} f(c)g(c) &= \lim_{h\to 0} \frac{f(c+h)g(c+h) - f(c)g(c)}{h} \\
&= \lim_{h\to 0} \frac{f(c+h)g(c+h) + f(c+h)g(c) - f(c+h)g(c) - f(c)g(c)}{h} \\
&= \lim_{h\to 0} f(c+h)\frac{g(c+h)-g(c)}{h} + g(c)\frac{f(c+h)-f(c)}{h} \\
&= \left[\lim_{h\to 0} f(c+h)\right]\left[\lim_{h\to 0} \frac{g(c+h)-g(c)}{h}\right] + g(c)\left[\lim_{h\to 0}\frac{f(c+h)-f(c)}{h}\right] \\
&= f(c)g'(c) + g(c)f'(c).
\end{aligned}$$

□

**Example 6.14**

Let $f$ be a differentiable function such that $f'(x) = 1/f(x)$. Compute $\frac{\mathrm{d}}{\mathrm{d}x}[f(x)]^2$.

*Solution.* Applying the product rule, we have

$$\frac{\mathrm{d}}{\mathrm{d}x}[f(x)]^2 = f'(x)f(x) + f'(x)f(x) = 2f'(x)f(x).$$

Now since we were told that $f'(x) = 1/f(x)$ we may substitute this to find that

$$\frac{\mathrm{d}}{\mathrm{d}x}[f(x)]^2 = 2f'(x)f(x) = 2\frac{f(x)}{f(x)} = 2.$$

■

**Exercise:** Find a differentiable function that satisfies $f'(x) = [f(x)]^{-1}$ as in Example 6.14.

**Example 6.15**

Let $f$ be differentiable and satisfy $f(1) = 1$ and $f'(1) = 2$. Compute $g'(1)$ where $g(x) = f(x)/x$.

*Solution.* We have already seen that $\frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{x} = -\frac{1}{x^2}$, so the product rule tells us that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}g(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left[f(x) \times \frac{1}{x}\right] &= f'(x)\left(\frac{1}{x}\right) + f(x)\left(\frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{x}\right) \\ &= \frac{f'(x)}{x} + f(x)\left(-\frac{1}{x^2}\right) \\ &= \frac{f'(x)}{x} - \frac{f(x)}{x^2}.\end{aligned}$$

If we now substitute $x = 1$ into this equation we find

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

■

In fact, there is no reason to limit ourselves to considering the product of only two functions. If $f, g$, and $h$ are all differentiable then

$$\frac{\mathrm{d}}{\mathrm{d}x}f(x)g(x)h(x) = f'(x)g(x)h(x) + f(x)g'(x)h(x) + f(x)g(x)h'(x).$$

The way to see this is to define a new function $n(x) = g(x)h(x)$ so that $n'(x) = g'(x)h(x)+g(x)h'(x)$ and $f(x)g(x)h(x) = f(x)n(x)$. Since the right-hand-side is a product of two functions, the product rule again gives us

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}f(x)n(x) = f'(x)n(x) + f(x)n'(x) &= f'(x)g(x)h(x) + f(x)\left[g'(x)h(x) + g(x)h'(x)\right] \\ &= f'(x)g(x)h(x) + f(x)g'(x)h(x) + f(x)g(x)h'(x)\end{aligned}$$

and this process is easily generalized to any number of functions.

**Theorem 6.16: The Quotient Rule**

If $f$ and $g$ are differentiable at $c$ and $g(c) \neq 0$ then $f/g$ is differentiable at $c$ and

$$\left.\frac{\mathrm{d}}{\mathrm{d}c}\right|_{x=c}\frac{f(x)}{g(x)} = \frac{f'(c)g(c) - f(c)g'(c)}{[g(c)]^2}.$$

**Example 6.17**

Compute the derivative of $f(x) = \dfrac{x^2 - 2x + 1}{x^4 + 4}$.

*Solution.* Let us write $g(x) = x^2 - 2x + 1$ and $h(x) = x^4 + 4$ so that $f = g/h$. We know that $g'(x) = 2x - 2$ and $h'(x) = 4x^3$ so

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\frac{x^2 - 2x + 1}{x^4 + 4} &= \frac{g'(x)h(x) - g(x)h'(x)}{[h(x)]^2} \\ &= \frac{(2x - 2)(x^4 + 4) - (4x^3)(x^2 - 2x + 1)}{[x^4 + 4]^2} \\ &= \frac{-2x^5 + 6x^4 - 4x^3 + 8x - 8}{[x^4 + 4]^2}\end{aligned}$$

■

**Example 6.18**

Confirm the computation of the derivative of $f(x)/x$ given in Example 6.15.

*Solution.* Applying the quotient rule to $f(x)/x$ we find that

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{f(x)}{x} = \frac{f'(x)x - \left(\frac{\mathrm{d}}{\mathrm{d}x}x\right)f(x)}{[x^2]} = \frac{xf'(x) - f(x)}{x^2}.$$

In Example 6.15 we found that

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{f(x)}{x} = \frac{f'(x)}{x} - \frac{f(x)}{x^2} = \frac{xf'(x) - f(x)}{x^2}$$

exactly as expected. ■

### 6.2.4 Higher Order Derivatives

Differentiating a function $f$ resulted in another function $f'$. If $f'$ is itself differentiable then we can apply the derivative again to find the second derivative $f''$. If $f''$ is differentiable, we can differentiate a third time to get $f'''$, and so on.

When using the prime notation becomes too cumbersome, we let $f^{(n)}(x)$ denote the $n^{th}$ derivative of $f$. These have important interpretations in both mathematics and science which we shall explore later. In Leibniz notation, we use the operator $\frac{\mathrm{d}}{\mathrm{d}x}$ to take subsequent derivatives, hence if $y = f(x)$ then the first derivative is $\frac{\mathrm{d}y}{\mathrm{d}x}$, while the second, third, and fourth derivatives are

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}^2y}{\mathrm{d}x^2}, \qquad \frac{\mathrm{d}}{\mathrm{d}x}\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\mathrm{d}^3y}{\mathrm{d}x^3}, \qquad \frac{\mathrm{d}}{\mathrm{d}x}\frac{\mathrm{d}^3y}{\mathrm{d}x^3} = \frac{\mathrm{d}^4y}{\mathrm{d}x^4}$$

respectively, with the pattern continuing *ad infinitum.* to

**Example 6.19**

Compute the second derivative of the function $f(x) = 1/x$. Determine a formula for the $n^{\text{th}}$ derivative $f^{(n)}(x)$.

*Solution.* Applying the quotient rule to $f$ quickly gives $f'(x) = -1/x^2$. To compute $f''$ we again apply the quotient rule and find that

$$f''(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(-\frac{1}{x^2}\right) = \frac{2}{x^3}.$$

Were we to continue on in this fashion, the higher order derivatives would be computed to be

$$f'''(x) = \frac{-6}{x^4}, \quad f^{(4)}(x) = \frac{24}{x^5}, \quad f^{(5)}(x) = -\frac{120}{x^6}, \ldots.$$

In general, the $n$-th derivative of $f$ is given by

$$f^{(n)}(x) = \frac{(-1)^n n!}{x^{n+1}}.$$ ■

**Example 6.20**

If $g(x) = x/e^x$, find an expression for $g^{(n)}(x)$.

*Solution.* Once we have learned the chain rule, this example will be much more easily computed by using the product rule on $g(x) = xe^{-x}$. For the moment though we must content ourselves with using the quotient rule. Since we would like to find a general expression for the $n^{th}$ derivative of $g(x)$, we will start by computing the first few derivatives and see if we can find a pattern. The first several derivatives are as follows:

$$\begin{aligned} f'(x) &= \frac{e^x - xe^x}{e^{2x}} = \frac{1-x}{e^x} \\ f''(x) &= \frac{-e^x - (1-x)e^x}{e^{2x}} = \frac{-1-1+x}{e^x} = \frac{-(2-x)}{e^x} \\ f'''(x) &= \frac{e^x - (x-2)e^x}{e^{2x}} = \frac{1-x+2}{e^x} = \frac{3-x}{e^x}. \end{aligned}$$

The pattern would suggest that in general,

$$g^{(n)}(x) = \frac{(-1)^{n-1}(n-x)}{e^x}. \qquad \blacksquare$$

## 6.3 Smoothness of Differentiable Functions

Differentiable functions are well behaved functions, in that their graphs appear to be smooth. This will be made more explicit when we discuss the different ways a function can fail to be differentiable, but we will start by discussing the relationship between continuity and differentiability.

### 6.3.1 Differentiable implies Continuous

**Proposition 6.21**

If $f$ is differentiable at $a$, then $f$ is continuous at $a$.

*Proof.* Let $f$ be differentiable at the point $a$, so that

$$\lim_{x\to a} \frac{f(x) - f(a)}{x-a} = L \text{ exists and is finite.} \tag{6.5}$$

To show continuity, we wish to show that

$$\lim_{x\to a} f(x) = f(a), \qquad \text{or equivalently} \qquad \lim_{x\to a} [f(x) - f(a)] = 0.$$

At this point we say to ourselves, "what part of the hypothesis have we failed to use?" Well, differentiability! Since our function is differentiable, we know something about the limit $(f(x) -$

$f(a))/(x-a)$ and now wish to say something about $f(x)-f(a)$. Multiplying and dividing by the quantity $(x-a)$ we find

$$\begin{aligned}
\lim_{x\to a}[f(x)-f(a)] &= \lim_{x\to a}\frac{f(x)-f(a)}{x-a}(x-a) \\
&= \left[\lim_{x\to a}\frac{f(x)-f(a)}{x-a}\right]\left[\lim_{x\to a} x-a\right] && \text{by the limit laws, since both limits exist} \\
&= L\times 0 = 0
\end{aligned}$$

as desired. □

It can be easy to confuse these two notions, so take a moment to get the direction of the implication correct. If a function is differentiable, then it is continuous. In the next section, we will see that the opposite direction is not true; namely, there are functions which are continuous but not differentiable.

### 6.3.2 Failures of Differentiability

There are three ways in which a function which fail to be differentiable.

1. The slope of the tangent line cannot be determined at a point.
2. The function may fail to be continuous.
3. The function may have a vertical tangent line.

The first condition can arise when each of the one-sided limits

$$\lim_{h\to 0^+}\frac{f(x+h)-f(x)}{h} \quad \text{and} \quad \lim_{h\to 0^-}\frac{f(x+h)-f(x)}{h}$$

exist, but are not equal. Since differentiability is defined in terms of the two sided limit, this would represent a failure of differentiability. This often presents in the graph of $f$ as a 'corner' or jagged edge.

Since all differentiable functions are continuous, if a function is not continuous then it certainly cannot be differentiable! Finally, a vertical tangent line has slope "infinity," meaning that the derivative diverges at that point.

**Example 6.22**

Show that the function

$$f(x)=\begin{cases} x^2 & x\geq -1 \\ x & x<-1 \end{cases}$$

is not differentiable at $x=-1$.

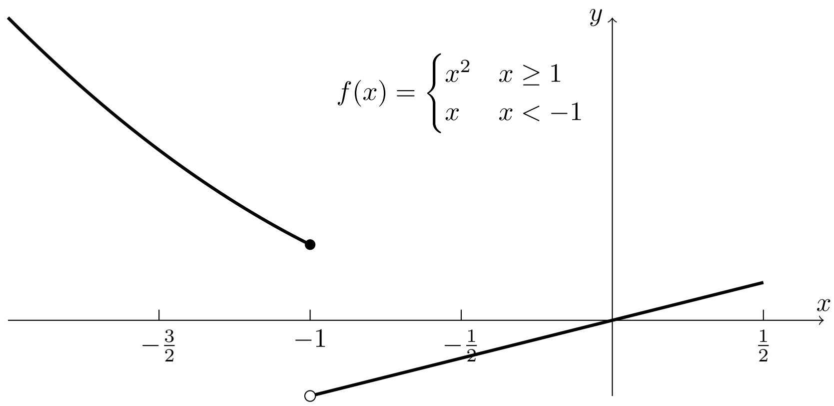



Figure 6.3: The graph of the function from Example 6.22. Since the function is not continuous at $x = -1$, it certainly cannot be differentiable there.

*Solution.* This function has a jump discontinuity at $x = -1$, and one cannot help but think this might affect its differentiability. Indeed, the one sided limits of the derivative yield

$$\lim_{h \to 0^+} \frac{f(1+h) - f(1)}{h} = \lim_{h \to 0} \frac{(1+h)^2 - 1}{h} = \lim_{h \to 0} \frac{h^2 + 2h}{h} = 2$$
$$\lim_{h \to 0^-} \frac{f(1+h) - f(1)}{h} = \lim_{h \to 0} \frac{(1+h) - 1}{h} = \lim_{h \to 0} \frac{h}{h} = 1$$

and these certainly do not agree. We conclude non-differentiability at $x = -1$ as required. ■

The discontinuity above feels (and is) contrived, so we offer a more enlightening example.

**Example 6.23**

Show that the function $f(x) = |x|$ is not differentiable at $x = 0$.

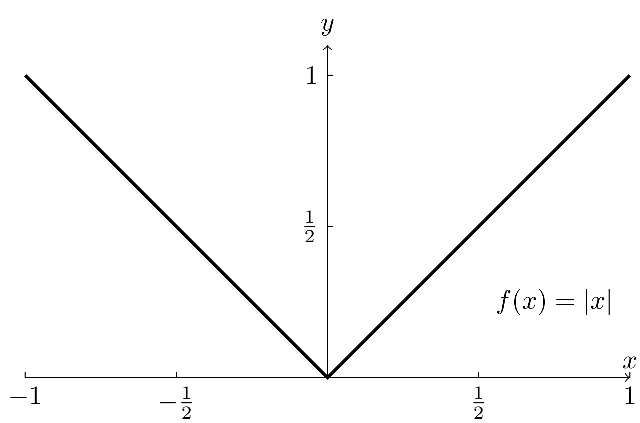



Figure 6.4: The graph of the function from Example 6.23. The corner at $x = 0$ keeps this function from being differentiable there.

*Solution.* Unlike Example 6.22, the function $f(x) = |x|$ is continuous. Nonetheless, the left- and right-sided limits disagree:

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

So continuous functions need not be differentiable! We saw in Proposition 6.21 that the converse is true: Every differentiable function is continuous. ■

The pathology in the previous example is due to the corner which occurs at $x = 0$. Since the derivative measures the rate of change of a function, the jarring change of $|x|$ at the origin is accountable for the lack of differentiability.

**Example 6.24**

The function $f(x) = \sqrt[3]{x}$ is not differentiable at $x = 0$.

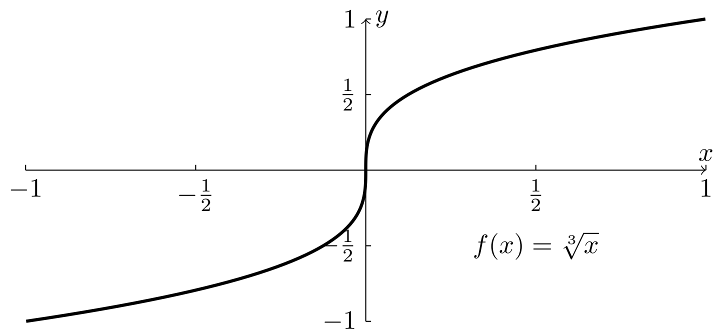



Figure 6.5: The graph of the function for Example 6.24. Note that the tangent line at $x = 0$ is vertical, and so has infinite slope.

*Solution.* This function does not have the abrupt changes exhibited in Example 6.23, but still fails to be differentiable at 0. Our one-sided limits give

$$\lim_{h\to 0^+} \frac{f(0+h)-f(0)}{h} = \lim_{h\to 0} \frac{\sqrt[3]{h}}{h} = \lim_{h\to 0} \frac{1}{h^{2/3}} = \infty.$$

This implies there is a sense in which the derivative exists, it is just that the tangent line has infinite slope; that is, the tangent line is vertical. As we would prefer to avoid having the incorporation of infinities into our formalism, we say that $f$ is not differentiable at 0 all the same. ■

## 6.4 Chains and Inverses

We have seen how to take the derivatives of sums, products, and quotients of functions. The only major operation left to look at is function composition. Interestingly, determining how to differentiate a composition will also give us access to the derivative of inverse functions.

### 6.4.1 The Chain Rule

Why are we interested in differentiating compositions? Let's say that you're a weatherperson and you want to determine the outside temperature over the course of the day. You realize that the temperature $H$ depends on the amount of sunshine $s$, and they are related through a function $H = f(s)$. This makes perfect sense, and you can even use calculus to determine $\frac{\mathrm{d}H}{\mathrm{d}s}$, the instantaneous rate of change of temperature with respect to sunshine.

But maybe you find it difficult to measure the amount of sunshine directly. Instead, you realize that sunshine itself is a function of time $t$; that is, you can write $s = g(t)$ for some function $g$. This allows you to determine $\frac{\mathrm{d}s}{\mathrm{d}t}$, the instantaneous rate of sunshine with respect to time.

The composition $H = f(s) = f(g(t))$ now tells you how the temperature depends on time. However, to differentiate this function you need to be able to differentiate the composition $f \circ g$. It seems like you should be able to do this; after all, you know how temperature changes with sunshine, and how sunshine changes with time:

$$\text{time} \longrightarrow \text{sunshine} \longrightarrow \text{temperature}.$$

Here, sunshine just acts as an intermediary for getting from time to temperature, and we can write $T = f(s) = f(g(t))$. This is our goal.

**Theorem 6.25: Chain Rule**

If $f$ and $g$ are functions such that $g$ is differentiable at $c$ and $f$ is differentiable at $g(c)$, then the composition $f \circ g$ is differentiable at $c$ and $(f \circ g)'(c) = f'(g(c))g'(c)$.
In Leibniz notation, if $w = f(y)$ and $y = g(x)$ then

$$\left.\frac{\mathrm{d}w}{\mathrm{d}x}\right|_a = \left.\frac{\mathrm{d}w}{\mathrm{d}y}\right|_{g(a)} \left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_a. \tag{6.6}$$

**Example 6.26**

Compute the derivative $\dfrac{\mathrm{d}}{\mathrm{d}x}(x^2 + 2x - 4)^{200}$.

*Solution.* If the only technique we know is the power rule, evaluation of this derivative would require us to expand the 200-fold product of $x^2 + 2x - 4$– what a mess! Instead, define the functions $f(x) = x^{200}$ and $g(x) = x^2 + 2x - 4$ so that $(x^2 + 2x - 4)^{200} = f(g(x))$. Using the fact that $f'(x) = 200x^{199}$ and $g'(x) = 2x + 2$ the chain rule gives us

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^2 + 2x - 4)^{200} = \frac{\mathrm{d}}{\mathrm{d}x} f(g(x)) = f'(g(x))g'(x) = 200(x^2 + 2x - 4)^{199}(2x + 2). \qquad \blacksquare$$

**Example 6.27**

Compute the derivative of $\sqrt{x + \sqrt{x}}$.

*Solution.* We need to realize $\sqrt{x+\sqrt{x}}$ as the composition of two functions. In particular, let $f(x)=\sqrt{x}$ and $g(x)=x+\sqrt{x}$ so that

$$f(g(x))=\sqrt{g(x)}=\sqrt{x+\sqrt{x}}.$$

Now we know that $f'(x)=1/(2\sqrt{x})$ and $g'(x)=1+1/(2\sqrt{x})$ so using the chain rule we have

$$\frac{\mathrm{d}}{\mathrm{d}x}\sqrt{x+\sqrt{x}}=\frac{\mathrm{d}}{\mathrm{d}x}f(g(x))=f'(g(x))g'(x)=\frac{1}{2\sqrt{g(x)}}\left(1+\frac{1}{2\sqrt{x}}\right)=\frac{2\sqrt{x}+1}{4\sqrt{x^2+x\sqrt{x}}}. \qquad \blacksquare$$

In Leibniz notation, set $y=f(g(x))$ and $u=g(x)$ so that $y=f(u)$. One could compute the derivative $\frac{\mathrm{d}y}{\mathrm{d}u}=f'(u)$ with no problem: This describes how the variable $y$ changes with respect to the variable $u$. However, if we want to know how $y$ changes with respect to the variable $x$, the chain rule is then written as

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x}$$

which makes it look surprisingly like fraction-cancellation.

As with the product rule, we may also extend the chain rule to three or more functions. For example, given the function $f(g(h(x)))$ let us temporarily define a new function $n(x)=g(h(x))$, whose derivative is computed by the chain rule to be $n'(x)=g'(h(x))h'(x)$. To compute our three-fold composition, we then have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}f(g(h(x)))&=\frac{\mathrm{d}}{\mathrm{d}x}f(n(x))\\&=f'(n(x))n'(x)\\&=f'(g(h(x)))g'(h(x))h'(x).\end{aligned}$$

To me, this looks a lot like a collection of Matryoshka dolls!

**Example 6.28**

Compute the derivative $\dfrac{\mathrm{d}}{\mathrm{d}x}e^{\sqrt{x^2+1}}$.

*Solution.* There are really three function compositions occurring in this question. Let $f(x)=e^x, g(x)=\sqrt{x}$, and $h(x)=x^2+1$ so that $e^{\sqrt{x^2+1}}=f(g(h(x)))$. The corresponding derivatives are $f'(x)=e^x, g'(x)=[2\sqrt{x}]^{-1}$ and $h'(x)=2x$, with the chain rule giving

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}e^{\sin^2(x)}&=\frac{\mathrm{d}}{\mathrm{d}x}f(g(h(x)))=f'(g(h(x)))g'(h(x))h'(x)\\&=\underbrace{e^{\sqrt{x^2+1}}}_{f'(g(h(x)))}\underbrace{\frac{1}{2\sqrt{x^2+1}}}_{g'(h(x))}\underbrace{2x}_{h'(x)}=\frac{xe^{\sqrt{x^2+1}}}{\sqrt{x^2+1}}.\end{aligned} \qquad \blacksquare$$

**Example 6.29**

If $y=e^{\pi w}$, $w=\sqrt{z}$ and $z=x^2+4x-1$, compute $\left.\dfrac{\mathrm{d}y}{\mathrm{d}x}\right|_{x=1}$.

*Solution.* There are two ways in which we may proceed. The first is to actually compose all of the functions and evaluate. In this case the composition yields $y = e^{\pi\sqrt{x^2+4x-1}}$ and using the chain rule we can its derivative to be

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\pi}{2}\frac{e^{\sqrt{x^2+4x-1}}(2x+4)}{\sqrt{x^2+4x-1}}.$$

Evaluating at $x = 1$ gives the value $3\pi/2$. The alternative technique is to use Leibniz notation: notice that at $x = 1$ we have $z = x^2 + 4x - 1 = 4$, $w = \sqrt{z} = 2$ and $y = e^{2\pi}$. This means that

$$\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{x=1} = \left[\frac{\mathrm{d}y}{\mathrm{d}w}\frac{\mathrm{d}w}{\mathrm{d}z}\frac{\mathrm{d}z}{\mathrm{d}x}\right]_{x=1} = \left.\frac{\mathrm{d}y}{\mathrm{d}w}\right|_{w=2}\left.\frac{\mathrm{d}w}{\mathrm{d}z}\right|_{z=4}\left.\frac{\mathrm{d}z}{\mathrm{d}w}\right|_{x=1}.$$

These may all be computed separately. Indeed, notice that $\frac{\mathrm{d}y}{\mathrm{d}w} = \pi e^{\pi w}$, $\frac{\mathrm{d}w}{\mathrm{d}z} = \frac{1}{2\sqrt{z}}$, and $\frac{\mathrm{d}z}{\mathrm{d}x} = 2x+4$. Hence

$$\left.\frac{\mathrm{d}y}{\mathrm{d}w}\right|_{w=2} = \pi e^{2\pi}, \qquad \left.\frac{\mathrm{d}w}{\mathrm{d}z}\right|_{z=4} = \frac{1}{4}, \qquad \left.\frac{\mathrm{d}z}{\mathrm{d}x}\right|_{x=1} = 6$$

which we may combine all together to find

$$\left.\frac{\mathrm{d}y}{\mathrm{d}w}\right|_{w=2}\left.\frac{\mathrm{d}w}{\mathrm{d}z}\right|_{z=4}\left.\frac{\mathrm{d}z}{\mathrm{d}w}\right|_{x=1} = \pi e^{2\pi} \times \frac{1}{4} \times 6 = \frac{3\pi}{2}e^{2\pi}.$$

■

**Example 6.30**

Let $f$ be a differentiable function. Show that for any positive integer $n$

$$\frac{\mathrm{d}}{\mathrm{d}x}f(x)^n = nf(x)^{n-1}f'(x).$$

*Solution.* By setting $g(x) = x^n$ we have that $f(x)^n = g(f(x))$. Furthermore, $g'(x) = nx^{n-1}$, so applying the chain rule we have

$$\frac{\mathrm{d}}{\mathrm{d}x}f(x)^n = g'(f(x))f'(x) = nf(x)^{n-1}f'(x)$$

as required. ■

**Example 6.31**

Denote by $f^{\circ n}(x)$ the $n$-fold composition of $f$. For example, $f^{\circ 3}(x) = f(f(f(x)))$. Assuming that $f(1) = 1$ and $f'(1) = 2$ find the derivative of $f^{\circ n}$ evaluated at $x = 1$.

*Solution.* This is just a repeated exercise of the chain rule requiring a small bit of trickery. First, we notice that since $f(1) = 1$ then no matter how many times we compose by $f$, we will always get 1. More explicitly, notice that $f(f(1)) = f(1) = 1$, and in general $f^{\circ n}(1) = 1$. For the sake of intuition, let us try this when $n = 3$. Notice in this case that

$$(f^{\circ 3})'(x) = f'(f(f(x))) \cdot f'(f(x)) \cdot f'(x)$$

so that

$$(f^{\circ 3})'(1) = f'(f(f(1))) \cdot f'(f(1) \cdot f'(1) = f'(1) \cdot f'(1) \cdot f'(1) = \left[f'(1)\right]^3 = 8.$$

In fact, precisely the same procedure will work for general $n$, and we get

$$(f^{\circ n})'(x) = f'(f^{\circ(n-1)}(x))f'(f^{\circ(n-2)}(x)) \cdots f'(f(x))f'(x),$$

where $f^{\circ 0} = x$. Hence we get

$$(f^{\circ n})'(1) = \underbrace{f'(1) \cdots f'(1)}_{n-\text{times}} = 2^n.$$ ■

Just as we were able to use the product rule to extend the power rule from $n \in \mathbb{N}$ to $n \in \mathbb{Z}$, we can use the product rule to tell us something about inverse functions (we will discuss this more in a subsequent section).

### 6.4.2 Derivatives of Inverse Functions

**Theorem 6.32: Inverse Function Theorem**

Let $f$ be differentiable at the point $c$ with $f'$ continuous at $c$. If $f'(c) \neq 0$, then there is an interval $I$ containing $c$ on which $f$ is invertible. Moreover, the inverse $f^{-1}$ is differentiable with continuous derivative, and satisfies the formula

$$\left(f^{-1}\right)'(x) = \frac{1}{f'(f^{-1}(x))}. \tag{6.7}$$

The majority of this theorem is beyond our abilities, but deriving Equation (6.7) is not too difficult. Let us begin by assuming that we have been given $f^{-1}$ and we know that it is differentiable. By definition of the inverse function we have $f(f^{-1}(x)) = x$ for all $x$ in the range of $f$. Differentiating both sides (applying the chain rule to the composition) we then get

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

We can solve for $(f^{-1})'(x)$ by re-arranging to get $(f^{-1})'(x) = \left[f(f^{-1}(x))\right]^{-1}$ as required. Again, this is an instance in which the derivation of the formula is so easy that it would be wasteful to memorize (6.7). Instead, we emphasize that the student should focus on the derivation itself.

**Proposition 6.33**

The derivative of $\ln(x)$ is

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

*Proof.* It is possible to do this using first principles, but the proof turns out to be difficult. Instead, we use the fact that $e^x$ and $\ln(x)$ are inverses, so that $e^{\ln(x)} = x$. Set $f(x) = e^x$ and $f^{-1}(x) = \ln(x)$, so that by the Inverse Function Theorem we have

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln(x) = (f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}$$
$$= \frac{1}{e^{\ln(x)}} = \frac{1}{x}. \qquad \square$$

Now that we know the derivative of $\ln(x)$, we can generalize the Power Rule, and differentiate general exponential functions $f(x) = a^x$:

**Theorem 6.34: Generalized Power Rule**

If $n$ is *any* real number, then $f(x) = x^n$ is differentiable for all $x > 0$, and moreover

$$\frac{\mathrm{d}}{\mathrm{d}x}x^n = nx^{n-1}.$$

*Proof.* Let $n$ be any real number. By properties of the logarithm and exponential, we can write

$$x^n = e^{\ln(x^n)} = e^{n\ln(x)}.$$

Setting $f(x) = e^x$ and $g(x) = n\ln(x)$, we can differentiate $e^{n\ln(x)} = f(g(x))$ using the chain rule:

$$\frac{\mathrm{d}}{\mathrm{d}x}x^n = \frac{\mathrm{d}}{\mathrm{d}x}e^{n\ln(x)} = f'(g(x))g'(x) = \frac{n}{x}e^{n\ln(x)} = \frac{n}{x}x^n = nx^{n-1},$$

which is precisely what we wanted to show. $\square$

**Theorem 6.35**

If $a > 0$ then the function $f(x) = a^x$ is differentiable for all $x$, and moreover

$$\frac{\mathrm{d}}{\mathrm{d}x}a^x = \ln(a)a^x.$$

*Proof.* Using the properties of exponents and logarithms, we can write

$$a^x = e^{\ln(a^x)} = e^{x\ln(a)}.$$

Using the chain rule, we set $f(x) = e^x$ and $g(x) = x\ln(a)$ so that $f'(x) = e^x$ and $g(x) = \ln(a)$ giving

$$\frac{\mathrm{d}}{\mathrm{d}x}a^x = \frac{\mathrm{d}}{\mathrm{d}x}e^{x\ln(a)} = \frac{\mathrm{d}}{\mathrm{d}x}f(g(x))$$
$$= f'(g(x))g'(x) = e^{x\ln(a)}\ln(a)$$
$$= \ln(a)a^x. \qquad \square$$

### 6.4.3 Logarithmic Differentiation

One of the great things about logarithms is there is a sense in which they decrease the complexity of an operation. For example, we often think of addition as being easier than multiplication, and multiplication being easier than exponents:

$$\ln(xy) = \ln(x) + \ln(y), \qquad \ln(x^y) = y\ln(x).$$

At the cost of introducing a logarithm, we are able to convert product to sums, and powers to products! Since the logarithm is not very hard to differentiate, this does not seem like such a terrible cost.

This idea in general is known as *logarithmic differentiation*. Where it can be particularly useful is when we have a product/quotient of many objects which are individually simple to differentiate, but which will become complicated when nested with the product rule. For example, given a collection of functions $f_1, \ldots, f_n$ and $g_1, \ldots, g_m$, notice that we can write

$$\ln\left[\frac{f_1(x)\cdots f_n(x)}{g_1(x)\cdots g_m(x)}\right] = \ln f_1(x) + \cdots + \ln f_n(x) - \ln g_1(x) - \cdots - \ln g_m(x).$$

Hence implicit differentiation yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\frac{f_1(x)\cdots f_n(x)}{g_1(x)\cdots g_m(x)} &= \frac{f_1(x)\cdots f_n(x)}{g_1(x)\cdots g_m(x)}\frac{\mathrm{d}}{\mathrm{d}x}\left[\ln f_1(x) + \cdots + \ln f_n(x) - \ln g_1(x) - \cdots - \ln g_m(x)\right] \\ &= \frac{f_1(x)\cdots f_n(x)}{g_1(x)\cdots g_m(x)}\left[\frac{f_1'(x)}{f_1(x)} + \cdots + \frac{f_n'(x)}{f_n(x)} - \frac{g_1'(x)}{g_1(x)} - \cdots - \frac{g_m'(x)}{g_m(x)}\right]\end{aligned}$$

**Example 6.36**

Compute the derivative of $f(x) = \dfrac{(x-1)^2(x^2+2)\sqrt{x}}{x^4+5}$.

*Solution.* This would be an absolute nightmare to compute using the quotient rule, so instead we use logarithmic differentiation. Taking the logarithm of both sides yields:

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

Differentiating implicitly gives

$$\begin{aligned}f'(x) &= f(x)\frac{\mathrm{d}}{\mathrm{d}x}\left[2\ln(x-1) + \ln(x^2+2) + \frac{1}{2}\ln(x) - \ln(x^4+5)\right] \\ &= \frac{(x-1)^2(x^3+2)\sqrt{x}}{x^4+5}\left[\frac{2}{x-1} + \frac{2x}{x^2+2} + \frac{1}{2x} - \frac{4x^3}{x^4+5}\right].\end{aligned}$$ ■

This takes care of converting products to sums, but now what about powers to products? Given two functions $f$ and $g$, let's try to differentiate $f(x)^{g(x)}$. The problem here is that neither the power rule, nor the rules for differentiating exponents can apply (in both of those cases, the function should only occur in the power or the base, but not both). To deal with this, we set $y = f(x)^{g(x)}$ so that $\ln y = g(x) \ln f(x)$. We can now differentiate implicitly:

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = g'(x) \ln f(x) + g(x)\frac{f'(x)}{f(x)}$$

which we may then solve for $\frac{\mathrm{d}y}{\mathrm{d}x}$ to get

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)^{g(x)} \left[ g'(x) \ln f(x) + g(x)\frac{f'(x)}{f(x)} \right]. \tag{6.8}$$

Like many of the formulae that we've derived, Equation (6.8) is not worth remembering on its own. Rather, what is important is remembering how this derivation was performed so that it can be repeated when necessary.

**Example 6.37**

Compute the derivative of $x^x$.

*Solution.* Setting $y = x^x$ we have $\ln y = x \ln x$. Differentiating implicitly we get

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \ln(x) + \frac{x}{x}$$

which we may solve for $\frac{\mathrm{d}y}{\mathrm{d}x}$ to get

$$\frac{\mathrm{d}y}{\mathrm{d}x} = x^x \left[ \ln(x) + 1 \right].$$

■

## 6.5 Exercises

6-1. For each function $f$ below, determine $f'(c)$ for the specified point $c$, using the limit definition

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

(a) $f(x) = 2x + 1, c = 1$

(b) $f(x) = 3\sqrt{x+1}, c = 3$

(c) $f(x) = \dfrac{4}{x+1}, c = 2$

(d) $f(x) = \dfrac{x}{1-x}, c = -3$

(e) $f(x) = |x|, c = 3$

(f) $f(x) = x|x|, c = -1$

(g) $f(x) = 3x^3 - x, c = 0$

(h) $f(x) = \dfrac{\sqrt{x}}{x-1}, c = 4$

6-2. For each function in Exercise 6-1, determine where the function is differentiable, and determine $f'(x)$ in general.

6-3. For each function $f$ and point $c$ in Exercise 6-1, determine the equation of the tangent line to $f$ through $c$.

6-4. Redo Exercise 6-1, but this time use the limit

$$f'(c) = \lim_{x\to c} \frac{f(x) - f(c)}{x - c}$$

to find the derivative.

6-5. Assume that $f, g, h$ are all differentiable. Using the product and quotient rules to find expressions for the following:

(a) $xf(x)$

(b) $\dfrac{f(x)}{x}$

(c) $\dfrac{x^2 f(x)}{x + g(x)}$

(d) $\dfrac{f(x)g(x)}{h(x)}$

(e) $\dfrac{f(x)}{g(x)} + \dfrac{g(x)}{h(x)}$

6-6. Using any method we have studied thus far, find the derivatives of the following functions:

(a) $x^2 e^x$

(b) $e^x + e^{-x} + xe^{-x}$

(c) $\sqrt{x}e^x$

(d) $\dfrac{e^x}{e^x + e^{-x}}$

(e) $xe^x + 2xe^{2x} + 3xe^{3x}$

6-7. Find the second derivative to each function given in Exercise 6-6.

6-8. Guess a general formula for the $n$th derivative of each function:

(a) $f(x) = e^{2x}$

(b) $g(x) = x^k$

(c) $h(x) = \dfrac{1}{1-x}$

6-9. Determine where each function fails to be differentiable:

(a) $f(x) = \sqrt[3]{x}$.

(b) $f(x) = \dfrac{3x^2 - 3x - 18}{2x^2 - 18}$

(c) $f(x) = x|x|$

(d) $f(x) = |x-1||x-2|$

(e) $f(x) = \sqrt[5]{x^2 - 1}$

6-10. Write down a function $f$ which satisfies each set of properties:

(a) $f$ is continuous and differentiable everywhere.

(b) $f$ is continuous everywhere but fails to be differentiable at $x = -2$.

(c) $f$ is defined at every real number, but fails to be differentiable at exactly two points.

6-11. Find the values of $a, b$ which makes $f$ everywhere differentiable.

(a) $f(x) = \begin{cases} x^2 & x \leq 2 \\ ax + b & x > 2 \end{cases}.$

(b) $f(x) = \begin{cases} ax^2 + bx + 6 & x \leq 1 \\ 2x^5 + 3x^4 + 4x^2 + 5x + 6 & x < 1 \end{cases}$

6-12. Differentiate each function:

(a) $f(x) = (x^2+x+1)^{200}$

(b) $f(x) = \sqrt{e^{2x}+x^2}$

(c) $f(x) = xe^{4x-5x^2}$

(d) $f(x) = \ln(3x^3+2x^2+x)$

(e) $f(x) = 2^{2^x}$

(f) $f(x) = x^2\ln(\sqrt{x^2+x})$

(g) $f(x) = \dfrac{e^{x^2}}{(x+1)^3}$

(h) $f(x) = \dfrac{\ln(x+\sqrt{x})}{1-e^x}$

6-13. Consider the following table:

| $x =$ | $-2$ | $-1$ | $0$ | $1$ | $2$ |
|---|---|---|---|---|---|
| $f(x)$ | 4 | 0 | 1 | 12 | $-2$ |
| $g(x)$ | 0 | 6 | 2 | 0 | 3 |
| $f'(x)$ | 2 | $-7$ | 5 | 1 | 1 |
| $g'(x)$ | $-1$ | 3 | $-12$ | 6 | 5 |

Determine the derivative of each function below:

(a) $f(g(x))$ at $x = 0$

(b) $f(f(x))$ at $x = -1$

(c) $g(f(x))$ at $x = 2$

(d) $g(g(x))$ at $x = -2$

6-14. Let $f(x) = x^4e^{\sqrt{x^2+3x}}$. Find the equation of the tangent line to $f$ at the point $x = 1$.

6-15. Find the equation of the tangent line to $f(x) = x + x^{2x}$ when $x = 1$.

6-16. Find the derivative of each of the following functions:

(a) $f(x) = x^{x+1}$

(b) $g(x) = x^x + x^{2x}$

(c) $h(x) = \ln(x)^x$

(d) $q(x) = \sqrt{x+x^x}$

6-17. By any means available to you, differentiate the function $f(x) = x^x x^{x-1} x^{x-2}$.

6-18. Let $f(x) = x^{(x^x)}$ and $g(x) = (x^x)^x$. If $h(x) = f(x) - g(x)$, what is $h'(2)$?

6-19. Define the functions

$$f(x) = \frac{e^x+e^{-x}}{2} \quad \text{and} \quad g(x) = \frac{e^x-e^{-x}}{2}.$$

(a) Compute $f'$ and express your answer using only the functions $f$, $g$, and their derivatives.

(b) Write down a formula for $f^{(n)}$ for all naturals $n \in \mathbb{N}$.

(c) Compute the derivative of the function $h(x) = f(x)^2 - g(x)^2$. By any means available to you, conclude that $h(x) = 1$ is a constant function.

6-20. Let $f$ and $g$ be the functions referenced in Exercise 6-19.

(a) Show that $f^{-1}(x) = \ln(x+\sqrt{x^2-1})$. *Hint:* It may help to let $z = e^x$, and then solve for $z$. This question is a bit tricky, so be sure to try it early and ask for help if you're uncertain what to do.

(b) Find $\dfrac{\mathrm{d}}{\mathrm{d}x}f^{-1}(x)$ directly using the expression given in part (a).

(c) Find $\frac{\mathrm{d}}{\mathrm{d}x} f^{-1}(x)$ using Equation (6.7) of the notes. Make sure you get the same answer as part (b).

# 7 Applications of Derivatives

Here we'll see how we can use derivatives to help us solve problems.

## 7.1 Implicit Differentiation

### 7.1.1 The Idea of Implicit Functions

The idea of implicit differentiation is that we may be given variables in which it is *implied* that those variables depend on other variables, though we may not be able to explicitly write that relationship down. For example, to this point we have typically seen examples where we might write $y = f(x)$, in which case it is clear that that changes in $x$ affect changes in $y$, as prescribed by the function $f$. We were then able to determine the rate of change of $y$ with respect to $x$ by computing $\frac{\mathrm{d}y}{\mathrm{d}x}$.

However, we can sometimes write relationships down without being able to solve for one variable as a function of the other; for example

$$\begin{aligned} x^2 + y^2 &= 25 \\ e^x + xy + y &= 1 \\ f(x,y) &= k \quad \text{for some constant } k \end{aligned}$$

We can convince ourselves that the variables $x$ and $y$ above depend on one another. For example, consider the equation of the circle $x^2 + y^2 = 25$. If we set $y = 5$ then $x$ is forced to be 0, while if we were to set $x = 3$ then $y$ would have to be one of $y = \pm 4$. However, there is no function which makes the relationship between $x$ and $y$ explicit, since as our above example indicates, a single $x$-value may correspond to two possible $y$-values, and hence the relationship is not one given by a function.[3]

As an alternative example, consider the volume $V$ of a cylinder as a function of its radius $r$ and height $h$:

$$V = \pi r^2 h.$$

This equation defines an explicit relationship between the three entities $V, r$, and $h$. However, what if our cylinder were made of metal, and we were told that as temperature $T$ increases, both the radius and the height increase, and conversely when temperature decreases, the radius and the

[3] You might be distressed at the fact that I have written $x^2 + y^2 = 25$ as an implicit equation, since certainly we could solve to find

$$y = \sqrt{25 - x^2},$$

but I claim that this actually not an explicit representation of this function. The reason is that, for example, both $(0,5)$ and $(0,-5)$ are solutions to this equation, but we are unable to recover $(0,-5)$ from the expression $y = \sqrt{25 - x^2}$.

height decrease? In that case, we are *implicitly* assuming that $r$ and $h$ are functions of temperature $T$. This means that $V$ is also implicitly a function of temperature, and we have

$$V(T) = \pi r(T)^2 h(T). \tag{7.1}$$

This implicit understanding that all the variables now depend on temperature allows us to determine how the volume of our cylinder is changing with temperature.

In both of the aforementioned cases, it seems as though we should still be able to discuss the rate of change of one variable with respect to another, even if we are unable to explicitly describe the relationship between the variables using a function. This leads us to a process known as implicit differentiation.

### 7.1.2 How Implicit Differentiation Works

Let's say that we know a variable $y$ implicitly depends upon another variable $x$. The idea of implicit differentiation is to differentiate as though the exact nature of the relationship were known. The best way to understand this is to see an example.

**Example 7.1**

Consider the equation of the circle centered at the origin with radius 1, $x^2+y^2 = 1$. Determine the rate of change of $y$ with respect to $x$.

*Solution.* Our goal is to compute $\frac{\mathrm{d}y}{\mathrm{d}x}$. We saw earlier that the equation of a circle is an implicit relationship as there is no function which describes how $y$ changes with respect to $x$ or vice versa. Nonetheless, we are going to differentiate the equation $x^2+y^2 = 1$, but we keep in mind always that we are assuming that $y$ is a function of $x$. To make this more clear, let's actually write $y = f(x)$, so that our equation of the circle is

$$1 = x^2 + y^2 = x^2 + f(x)^2.$$

Now differentiating both sides, we have

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}x}\left(x^2 + f(x)^2\right) \\ &= 2x + 2f(x)f'(x). \end{aligned}$$

Remember that we are trying to solve for $\frac{\mathrm{d}y}{\mathrm{d}x}$, which under our choice of $y = f(x)$ is just $\frac{\mathrm{d}y}{\mathrm{d}x} = f'(x)$, hence

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f'(x) = -\frac{x}{y}. \tag{7.2}$$

**Remark 7.2** Many people do not like to use this $y = f(x)$ notation, and instead will just write down

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(x^2 + y^2\right) = 2x + 2y\frac{\mathrm{d}y}{\mathrm{d}x}.$$

This is acceptable and if you are comfortable using it, then you should feel free to do that. However, writing this down often hides what is really happening, so we have used the $y = f(x)$ notation just to be clear.

On the other hand, let me write $\tilde{y} = \sqrt{1 - x^2}$, where the fact that I have used the tilde is to indicate that $\tilde{y}$ is not actually the same thing as $y$. When we differentiate we get

$$\frac{\mathrm{d}\tilde{y}}{\mathrm{d}x} = \frac{-x}{\sqrt{1 - x^2}} = -\frac{x}{\tilde{y}}. \tag{7.3}$$

The inquisitive student may realize looks very similar to (7.2), but I claim is not quite the same.

Indeed, let's try to find the slope of the tangent line to the circle at the point $x = 1/\sqrt{2}$. Notice on the circle that there are two possible $y$ values, corresponding to $y_+ = +1/\sqrt{2}$ and $y_- = -1/\sqrt{2}$. Using (7.2) we find that the slope of the tangent lines at $y_\pm$ are

$$\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)} = -\frac{1/\sqrt{2}}{1/\sqrt{2}} = -1 \quad \text{and} \quad \left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)} = -\frac{1/\sqrt{2}}{-1/\sqrt{2}} = 1$$

which is in fact what we would expect. On the other hand, using (7.3) we find that at $x = \frac{1}{\sqrt{2}}$ there is only one possible $\tilde{y}$ value, corresponding to $\tilde{y} = \frac{1}{\sqrt{2}}$ in which case the slope of the tangent line is the same as that found above, namely

$$\left.\frac{\mathrm{d}\tilde{y}}{\mathrm{d}x}\right|_{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)} = -\frac{1/\sqrt{2}}{1/\sqrt{2}} = -1.$$

We have in fact lost the other tangent line! This is because when we took the square root of $y^2 = 1 - x^2$ we needed to make a choice as to whether to take the positive or negative root. In doing so, we actually lost information. ■

**Example 7.3**

Consider the volume of a cylinder as a function of temperature, as given in (7.1). Determine the rate of change of $V$ with respect to temperature, written in terms of how $r$ and $h$ vary with respect to temperature.

*Solution.* We already know that $V = \pi r^2 h$. Although (7.1) has the temperature dependence written in, we will ignore it for this exercise to show the student how the notation is typically conveyed. Differentiating, we get

$$\frac{\mathrm{d}V}{\mathrm{d}T} = \frac{\mathrm{d}}{\mathrm{d}T}\left[\pi r^2 h\right] = \pi\left[2r\frac{\mathrm{d}r}{\mathrm{d}T}h + r^2\frac{\mathrm{d}h}{\mathrm{d}T}\right]. \qquad ■$$

The power of implicit differentiation can be even greater if one is given a transcendental equation such as $e^x + x\ln(y) + y = 1$, wherein it is impossible to isolate and solve for $y$. In such instances, one is indeed *forced* to use implicit differentiation.

**Example 7.4**

Compute the derivative $\frac{\mathrm{d}y}{\mathrm{d}x}$ of $y$ in the equation $e^x + x\ln(y) + y = 1$.

*Solution.* Keeping in mind that $y$ is a function of $x$, we apply $\frac{\mathrm{d}}{\mathrm{d}x}$ to both sides of our equation to find:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\left(e^x + x\ln(y) + y\right) &= \frac{\mathrm{d}}{\mathrm{d}x}1 \\ e^x + \ln(y) + \frac{x}{y}\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{\mathrm{d}y}{\mathrm{d}x} &= 0 \\ \frac{-y[e^x + \ln(y)]}{x+y} &= \frac{\mathrm{d}y}{\mathrm{d}x}.\end{aligned}$$

■

## 7.2 Rates of Change

We saw at the beginning of this section how derivatives can be used to deduce the *instantaneous rate of change* of one quantity with respect to another. The power of this is that it allows us to take boring, static relationships, and differentiate them to discover the dynamic interplay between variables.

**Example 7.5**

The volume of a sphere of radius $r$ is $V = \frac{4}{3}\pi r^3$. Determine how the volume changes as $r$ is allowed to vary.

*Solution.* All we need to do is differentiate the given expression with respect to $r$ to find that

$$\frac{\mathrm{d}V}{\mathrm{d}r} = 4\pi r^2.$$

This says that the rate of change of volume $V$ with respect to the radius $r$ is $4\pi r^2$. As an example, this means that doubling the radius of a sphere will quadruple its volume, while tripling the radius will increase the volume 9-fold. ■

One of the more utilized relationships is that of position, velocity, acceleration, jerk, etc. If $p(t)$ describes the position of an object with respect to time $t$, then $p'(t)$ is its velocity and $p''(t)$ is its acceleration with respect to time. This can be used to model physical situations which can then be solved by mathematical methods:

### 7.2.1 Economics

In economics, the word "marginal" is synonymous with "instantaneous rate of change." For example, if $c = f(q)$ is cost function relative to the quantity $q$ of an item produced, the *marginal cost* is $\frac{\mathrm{d}c}{\mathrm{d}q}$. Another example is if $C = g(Y)$ describes consumption as a function of income, then $\frac{\mathrm{d}C}{\mathrm{d}Y}$ is the *marginal propensity to consume.*

**Example 7.6**

Suppose that your saving as a function of your income is determined to be

$$S = 100 \ln \left( \frac{3}{2 + e^{-Y/10}} \right),$$

where both $S$ and $Y$ are measured in thousands of dollars. Determine your marginal propensity to save when your income is $\$30,000$/year.

*Solution.* An income of $\$30,000$/year corresponds to $Y = 30$. Differentiating $S$ and evaluating at $Y = 30$ we get

$$\left.\frac{\mathrm{d}S}{\mathrm{d}Y}\right|_{Y=30} = \left[ 100 \frac{2 + e^{-Y/10}}{3} \frac{3e^{-Y/10}}{10(2 + e^{-Y/10})^2} \right] = \frac{10}{2e^3 + 1} \approx 0.243.$$

This means that at the $\$30,000$ point, your spending habits are such that you save approximately 24 cents for each dollar you spend. ■

*Elasticity of demand* measures how demand for a product will change with an increase in price. More specifically, it is the ratio

$$\text{Elasticity of demand} = \frac{\text{percent change in quantity}}{\text{percent change in price}}.$$

Suppose then that $p = f(q)$ is price as a function of quantity, where $f$ is some differentiable function. Suppose that we increase the number of units from $q$ to $q + h$, so that

$$\text{percent change in quantity over } h \text{ units} = \frac{(q + h) - q}{h} = \frac{h}{q}.$$

The same holds true for price, with

$$\text{percent change in price over } h \text{ units} = \frac{f(q + h) - f(q)}{f(q)}.$$

When dividing these quantities and taking a limit as $h \to 0$, we get

$$\text{Point elasticity of demand} = \lim_{h \to 0} \frac{h/q}{(f(q + h) - f(q))/h} = \lim_{h \to 0} \frac{\frac{f(q)}{q}}{\frac{f(q+h) - f(q)}{h}} = \frac{\frac{p}{q}}{\frac{\mathrm{d}p}{\mathrm{d}q}}.$$

**Example 7.7**

Suppose the demand equation is specified by $q^2(1 + p)^2 = p$. Determine the point elasticity of demand when $p = 4$.

*Solution.* When $p = 4$ the equation becomes $25q^2 = 4$ so that $q = \pm 2/5$. A negative quantity is meaningless, so we discard it and have $q = 2/5$. Differentiating implicitly,

$$2q(1+p)^2 + 2q^2(1+p)\frac{\mathrm{d}p}{\mathrm{d}q} = \frac{\mathrm{d}p}{\mathrm{d}q} \quad \Rightarrow \quad \frac{\mathrm{d}p}{\mathrm{d}q} = \frac{2q(1+p)^2}{1-2q^2(1+p)},$$

into which we can substitute $q = 2/5$ and $p = 4$ to get $\left.\frac{\mathrm{d}p}{\mathrm{d}q}\right|_{p=4} = -100/3$. The point elasticity of demand is thus

$$\frac{p}{q}\left[\frac{\mathrm{d}p}{\mathrm{d}q}\right]^{-1} = \frac{10}{-100/3} = -\frac{3}{10}.$$ ■

### 7.2.2 Exponential Growth

Let $P(t)$ be an object which grows in proportion to its size. For example, consider a species of bacteria in which each bacterium splits and doubles after a period of 5-minutes. A colony of 100 bacteria will grow to 200 after 5-minutes, resulting in a growth of 100 bacteria, while a colony of 1000 bacteria will double to 2000 bacteria in 5 minutes, resulting in a growth of 1000 new bacteria! It should be clear then that the more bacteria present in the colony, the faster the colony will grow.

If we think about this example in more detail, let us say we start with a colony of a single bacteria. The colony size (specified over 5-minute intervals) will look like

$$1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096, 8192, 16384, 32768, 65536, \ldots$$

In general, this growth appears to be exponential, corresponding to a discrete function $2^x$.

Modelling exponential growth is often simplified by assuming continuous rather than discrete growth, and it turns out that this is actually one of the motivating example for defining Euler's number $e$. Perhaps the easiest way to see this is in the language of investment. Assume that you are given an initial investment $I_0$ which grows with an interest rate $r$. If we compound the interest annually (once per year) then after one year we have $I_0(1+r)$ dollars. If we compound the interest semi-annually (twice per year) then we take half the interest rate $r/2$ and compound twice to get

$$I_0\left(1+\frac{r}{2}\right)\left(1+\frac{r}{2}\right) = I_0\left(1+\frac{r}{2}\right)^2.$$

Similarly, if we compounded the interest $n$ times in a year, we would be left with the equation $I_0\left(1+\frac{r}{n}\right)^n$. The idea of continuously compounding interest will then occur as we let $n \to \infty$. Namely, the amount of money earned after one year will be

$$I_0 \lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n.$$

Now what happens if we make the change of variable $x = \frac{1}{n}$ so that as $n \to \infty$ we have $x \to 0$. Our equation then becomes

$$I_0 \lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n = I_0 \lim_{x\to 0}(1+rx)^{1/x} = I_0 e^r;$$

that is, the number $e$ represents the proportion by which your money will increase if compounded continuously over the course of a year. If we had started with \$1 then after one year of continuous interest at 1% we would finish the year with \$$e$. This is the sense to which $e$ is *natural*, it represents continuous growth.

Now if we return to our bacteria example, let us assume that our bacteria population grows at a continuous rate which is proportional to its population. That is, there is some constant of proportionality $k$ such that $\frac{\mathrm{d}P}{\mathrm{d}t} = kP$. What kind of function satisfies this differential equation? A bit of inspection actually reveals that if $I_0$ is the initial population, $P(t) = I_0 e^{kt}$ satisfies the equation, since

$$\frac{\mathrm{d}P}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} I_0 e^{kt} = I_0 k e^{kt} = kP(t).$$

Hence any quantity which grows (or shrinks) continuously and proportional to its population size is modelled with an exponential function.

**Example 7.8**

If an original investment of \$100 is invested at a rate of 6% and compounded continuously, how long will it take for the investment to triple in size?

*Solution.* Our model is given by $I(t) = 100e^{0.06t}$ and we would like to find the $t$ such that $I(t) = 300$ (since 300 is triple the number 100). Indeed, we may solve to find that $300 = 100e^{0.6t}$ implies

$$3 = e^{0.06t} \quad \Rightarrow \quad 0.06t = \ln(3) \quad \Rightarrow \quad t = \frac{\ln(3)}{0.06} \approx 18.3 \text{ years.} \qquad \blacksquare$$

**Example 7.9**

Assume that a culture of bacteria has an initial population of 100 bacteria, which becomes 350 bacteria after 10-hours. Determine the population of the bacteria after 2-days.

*Solution.* We know that our growth curve is modelled by the formula $P(t) = 100e^{kt}$, for $t$ measured in hours. We are also told that

$$P(10) = 100e^{10k} = 350,$$

which will allow us to solve for the growth rate $k$. Indeed,

$$100e^{10k} = 350 \quad \Rightarrow \quad e^{10k} = 3.5 \quad \Rightarrow \quad 10k = \ln(3.5) \quad \Rightarrow \quad k = \frac{\ln(3.5)}{10}.$$

To be consistent with our choice of units, we note that two days is 48-hours, so the population after two days is

$$P(48) = 100e^{48k} = 100e^{48\frac{\ln(3.5)}{10}} = 100e^{\ln(3.5^{4.8})} = 100 \cdot 3.5^{4.8} \approx 40881 \text{ bacteria.} \qquad \blacksquare$$

**Example 7.10**

Caffeine in the blood stream has a biological half-life of 5-hours. If I drink a venti café Americano from Starbucks, consisting of 300mg of caffeine, at 8am in the morning. How much caffeine will be in my system by midnight?

*Solution.* Let us that $t = 0$ corresponds to 8am, so that midnight is $t = 16$. The amount of caffeine is modelled by the equation $c(t) = 300e^{kt}$ and we know that $c(5) = 300e^{5k} = 150$ (since 150 is half of 300). We may solve for $k$ to find that $k = -\frac{\ln(2)}{30}$. Hence at time $t = 16$ we have

$$c(16) = 300e^{-16\frac{\ln(2)}{5}} = 300 \cdot 2^{-16/5} \approx 32.64 \text{ mg}. \qquad \blacksquare$$

One can show that if the half life of a substance is a time $t_0$, then the growth/decay rate $k$ is always given by $k = -\frac{\ln(2)}{t_0}$.

**Logistic Growth:** While this model of exponential growth can be quite useful for modelling short term growth, it quickly becomes unrealistic. Exponential functions grow incredibly quickly: in fact, if a species of bacteria grow at the same rate as in Example 7.9, then a single bacterium would grow into a colony with the same mass as the entire Earth in about one month!

Our exponential model breaks down in the long term as it fails to incorporate things like competition with other bacteria, or the limited number of resources available to the bacteria. If we know *a priori* that the bacteria only have enough resources to sustain a certain size colony, we can adapt our model to take that into consideration. The maximum number $M$ of the species is called the *carrying capacity*, and changes our model to be

$$\frac{\mathrm{d}P}{\mathrm{d}t} = kP\left(1 - \frac{P}{M}\right). \tag{7.4}$$

Saying that there is no maximum species is akin to setting "$M = \infty$," in which case (7.4) just becomes the usual $\frac{\mathrm{d}P}{\mathrm{d}t} = kP$. On the other hand, in the limit as $P \to M^-$ we get

$$\lim_{P\to M^-} \frac{\mathrm{d}P}{\mathrm{d}t} = \lim_{P\to M^-}\left[kP\left(1 - \frac{P}{M}\right)\right] = 0,$$

showing that our growth slows as our population approaches carrying capacity. The solution to (7.4) is not easy to guess, but is given by

$$P(t) = \frac{P_0 M e^{kt}}{P_0(e^{kt} - 1) + M}. \tag{7.5}$$

The asymptotic nature of this function can be seen in Figure 7.1.

**Exercise:** Show that (7.5) satisfies (7.4).

**Example 7.11**

As a biologist for a famous national park, your job is to ensure that the park maintains a stable ecology. In particular, you are tracking a group of red squirrels, which you know reproduce in an exponential fashion with growth coefficient $k = 0.32$, over a period of one year. The park has a theoretical capacity of 15000 squirrels, but you know that if the population reaches 80% of this number, it will begin to destabilize the surrounding populations. If the current population sits at 1300 squirrels, how long do you have before the squirrel population begins to pose a problem?

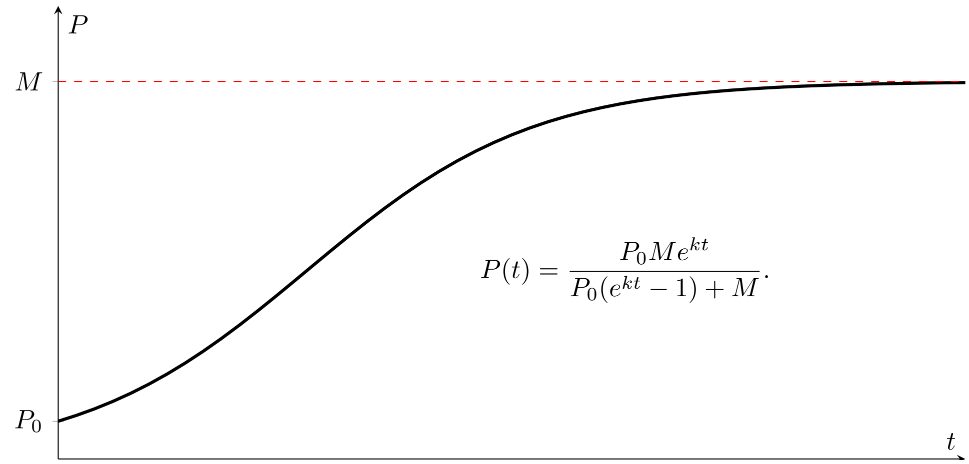



Figure 7.1: Logistic growth incorporates a carrying capacity $M$ so that the population cannot run off to infinity.

*Solution.* We are given a growth rate of $k = 0.32$, an initial population of $P_0 = 1300$, and a carrying capacity of $M = 15000$. The squirrels will become troublesome when they reach 80% of our carrying capacity, which is 12000 squirrels. Solving

$$12000 = \frac{1300 \times 15000 e^{0.32t}}{1300(e^{0.32t}-1)+15000}$$

we get $t = 11.7$ years. ■

## 7.3 Derivatives and the Shape of a Graph

### 7.3.1 First Derivative Information

Knowing something about the derivative of a function can often lead us to insights about what it's graph looks like.

**Theorem 7.12**

Let $f$ be a differentiable function with domain $(a, b)$. If $f'(x) = 0$ for all $a < x < b$ then $f$ is a constant function.

**Definition 7.13**

We say that a function $f$ is *non-decreasing* if whenever $a < b$ then $f(a) \leq f(b)$, and *strictly increasing* if whenever $a < b$ then $f(a) < f(b)$. In the same vein, we say that $f$ is *non-increasing* if whenever $a < b$ then $f(a) \geq f(b)$, and *strictly decreasing* if whenever $a < b$ then $f(a) > f(b)$.

**Theorem 7.14**

If $f$ is a differentiable function with domain $(a,b)$ and $f'(x) > 0$ for all $a < x < b$ then $f$ is strictly increasing.

The above theorem can also be used to determine where a function $f$ is non-increasing and non-decreasing, by determining the intervals on which $f'(x) > 0$ and $f'(x) < 0$ respectively.

**Example 7.15**

Show that the function $f(x) = e^{x^3}$ is everywhere strictly increasing.

*Solution.* Differentiating $f$ we get

$$f'(x) = 3x^2 e^{x^3}.$$

The exponential is always positive, as is $3x^2$, so $f'(x) > 0$ for all $x$. Hence by Theorem 7.14 we know that $f$ is everywhere strictly increasing. ■

If a function is going to change from increasing to decreasing, its derivative will pass through zero or a singularity. For this reason, we define the following:

**Definition 7.16**

If $f$ is a differentiable function and $c$ is in the domain of $f$, then we say that $c$ is a *critical point* of $f$ if $f'(c) = 0$ or $f'(c)$ does not exist.

Critical points therefore represent the possible points where a function could change from increasing to decreasing.

**Example 7.17**

Compute the critical points of the function $f(x) = \ln(x^2 - 1)$.

*Solution.* Differentiating yields

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

which is zero when $x = 0$ and does not exist when $x = \pm 1$. Hence the critical points of $f$ are $x = -1, 0, 1$. ■

**Example 7.18**

Determine the intervals on which the function $f(x) = 2x^3 + 3x^2 - 12x + 15$ is increasing and decreasing.

*Solution.* Differentiating our function gives

$$f'(x) = 6x^2 + 6x - 12 = 6(x^2 + x - 2) = 6(x + 2)(x - 1).$$

The derivative is zero at $x = -2$ and $x = 1$, so we check the surrounding intervals to determine the sign of $f'(x)$:

| | $x < -2$ | $-2 < x < 1$ | $1 < x$ |
|---|---|---|---|
| $x+2$ | $-$ | $+$ | $+$ |
| $x-1$ | $-$ | $-$ | $+$ |
| $f'(x)$ | $+$ | $-$ | $+$ |

We conclude that our function is strictly increasing on $(-\infty, -2)$ and $(1, \infty)$. ■

### 7.3.2 Second Derivative Information

**Definition 7.19**

A function $f$ is defined to be *concave down* on an interval $[a, b]$ if for every $c \in [a, b]$ the graph of the function $f$ as restricted to $[a, b]$ lies beneath the tangent line to $f$ at the point $c$. Similarly, $f$ is said to be *concave up* on $[a, b]$ if for every $c \in [a, b]$ the graph of the function $f$ restricted to $[a, b]$ lies above the tangent line to $f$ at the point $c$.

This is a rather cumbersome definition: to determine whether a function is concave up/down one would need to find the equation of a tangent line at each point in an interval and then manipulate complex inequalities to show that the graph of the function lies below/above that tangent line. As with most concepts in mathematics, the introductory idea is complicated and cumbersome, but often gives way to a much more simple characterization with the use of more sophisticated tools.

**Definition 7.20**

We say that a point $c$ is an *inflection point* of $f$ if $f$ changes from being convex up to convex down (or vice-versa) at the point $c$.

**Proposition 7.21**

Let $f$ be a function and $c$ be some point in the domain of $f$.

1. If $f''(c) > 0$ then $f$ is concave up at $c$,
2. If $f''(c) < 0$ then $f$ is concave down at $c$,
3. If $f''(c) = 0$ then let $k$ be the smallest positive integer such that $f^{(k)}(c) \neq 0$. If $k$ is odd then $c$ is an inflection point. Otherwise, let $k$ be even.
   (a) If $f^{(k)}(c) > 0$ then $f$ is concave up at $c$,
   (b) If $f^{(k)}(c) < 0$ then $f$ is concave down at $c$.

Proposition 7.21 thus tells us that points where $f''(x) = 0$ are candidates for inflection points.

**Example 7.22**

Let $f(x) = \ln(x^2+1) - x$. Determine and classify the critical points of $f$, find where $f$ is increasing/decreasing, and determine the intervals of concavity for $f$.

*Solution.* To solve for the critical points as well as increasing/decreasing, we must compute the first derivative:

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

The only critical point of this function thus corresponds to $x = 1$ (since $f'(1) = 0$). Furthermore, we can see that $f'(x) \leq 0$ for all $x$, since both $(x-1)^2$ and $x^2+1$ are always non-negative, so $f(x)$ is decreasing on all of $\mathbb{R}$. To classify the critical point and compute concavity, we use the second derivative:

$$f''(x) = -\left[\frac{2(x-1)(x^2+1) - 2x(x-1)^2}{(x^2+1)^2}\right] = -\frac{2(x-1)(x+1)}{(x^2+1)^2}.$$

Let us check concavity first, which corresponds to determining the sign of $f''(x)$. Since the denominator is always positive, this reduces to determining the sign of $-(x-1)(x+1)$, which we do with the following table:

| | $x < -1$ | $-1 < x < 1$ | $x > 1$ |
|---|---|---|---|
| $x+1$ | $-$ | $+$ | $+$ |
| $x-1$ | $-$ | $-$ | $+$ |
| $-(x-1)(x+1)$ | $-$ | $+$ | $-$ |

Hence the function is concave down on $(-\infty, -1) \cup (1, \infty)$ and concave up on $(-1, 1)$. This tells us that $-1$ and $1$ are the inflection points, which is actually sufficient to tell us that the critical point $x = 1$ is neither a max nor a min. Of course, we could also compute the third derivative

$$f^{(3)}(x) = -2\left[\frac{2x(x^2+1)^2 - 4x(x^2+1)(x^2-1)}{(x^2+1)^4}\right] = \frac{4x(x^2-3)}{(x^2+1)^3}$$

for which $f^{(3)}(1) \neq 0$. Our concavity criterion hence corroborates the fact that $x = 1$ is an inflection point. ■

**Example 7.23**

Consider the function $f(x) = \dfrac{1+x}{1+x^2}$. Determine all inflection points of this curve.

*Solution.* We plug along until we get to the second derivative of $f$. The first derivative is given by

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

while the second derivative is

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

The inflection points come from finding the zeroes of the numerator. It is not too hard to see that $x = 1$ is a zero, so we can factor this to get

$$x^3 + 3x^2 - 3x - 1 = (x-1)(x^2 + 4x + 1)$$

and this second term has roots $x = -2 \pm \sqrt{3}$. It is not too hard to convince ourselves that these are inflection points (an easy argument comes from the fact that there are three distinct linear roots with no multiplicities, hence the sign of the third derivative will change after passing any one root, implying each root is a proper inflection point). ■

## 7.4 Maxima and Minima

This section describes one of the most important applications of calculus: the ability to find necessary conditions for a point to be an extreme point of a function. Solving such problems is of exceptional importance. For example, all of classical mechanics works on the principle of *least* action, which says that a system will always try to minimize the difference between potential and kinetic energy. Relativity theory implies that gravity operates by moving particles along geodesics: paths of *minimal* length. When we sent people to the moon it was important to do it in the *quickest* amount of time while using the *least amount of fuel*. In business, we often like to *maximize* profits while *minimizing* waste. So how does calculus give us the ability to find extreme points? In order to make sense of this, we should first define what it means to be a max/min!

**Definition 7.24**

Let $f$ be a function with domain $D$. We say that $c$ is an *absolute maximum* if $f(x) \leq f(c)$ for all $x \in D$, and a *local maximum* if $f(x) \leq f(c)$ for all $x$ in a interval around $c$. Similarly, we say that $c$ is an *absolute minimum* if $f(x) \geq f(c)$ for all $x \in D$, and a local minimum if $f(x) \geq f(c)$ for all $x$ in a interval around $c$.

Specifying the domain is very important. For example, the function $f(x) = x$ has no global/local maximum/minimum on $\mathbb{R}$ or $(0,1)$, but over the interval $[0,1]$ it has a minimum at $x = 0$ and a maximum at $x = 1$. The following useful theorem tells us that intervals of the form $[a,b]$ ensure that maxima and minima always exist:

**Theorem 7.25: Extreme Value Theorem**

If $f$ is a continuous function on a closed interval $[a,b]$, then $f$ will attain its maximum and minimum on $[a,b]$.

**Example 7.26**

Determine on which of the following intervals the function $f(x) = 1/x$ is guaranteed to attain its global maximum and minimum.

$$[-1,1], \qquad (0,1), \qquad [1,2].$$

*Solution.* According to the Extreme Value Theorem, we can be guaranteed that $1/x$ attains its max and min if it is continuous on a closed interval. The first interval $[-1,1]$ is closed but $1/x$

is not continuous at 0 which is a point in $[-1,1]$. Hence we cannot guarantee that $1/x$ attains a max and min on $[-1,1]$, though note that it does attain global minima at $x = \pm 1$. Similarly, the interval $(0,1)$ is not closed so we cannot guarantee that the maximum or minimum is attained. In fact, there is no max or min of $1/x$ in $(0,1)$. Finally, $1/x$ *is* continuous on $[1,2]$ and so by the Extreme Value Theorem the max and min are attained. ■

This is all fine and dandy, but this is an existential theorem, meaning that it tells us when extrema exist but fails to provide any information on how to find them. In fact, we notice that the Extreme Value Theorem only requires the concept of continuity and so does not really fall within the regime of calculus. The real power of calculus is to provide a necessary condition for a point to be an extreme point.

**Theorem 7.27**

Let $f$ be a differentiable function on the interval $[a,b]$. If $c \in [a,b]$ is a local max or min of the function $f$, then it is necessarily a critical point of $f$.

The idea is that if a max/min occurs on the interior of the domain, then the function must curve back on itself. Think again of the function $f(x) = x^2$ which we noticed had a minimum at $x = 0$. We could ensure it was a local min (and in fact a global on) because the function decreased until it hit $x = 0$, then started to increase again. This means that at some point, the slope $f'$ of the function must have been zero. Similarly, if the max/min occurs at the endpoint of a domain then the derivative there did not exist. The idea behind why maxima and minima on the interior of the domain correspond to critical points gives us the following test for maximality/minimality:

**Theorem 7.28: First Derivative Test**

Suppose $f$ is a differentiable function with domain $[a,b]$ and critical point $c \in (a,b)$.

1. If there exists $\delta > 0$ such that $f'(x) > 0$ on $(c-\delta, c)$ and $f'(x) < 0$ on $(c, c+\delta)$ then $c$ is a maximum.
2. If there exists $\delta > 0$ such that $f'(x) < 0$ on $(c-\delta, c)$ and $f'(x) > 0$ on $(c, c+\delta)$ then $c$ is a minimum.

**Example 7.29**

Find the (local) maxima and minima of the function $f(x) = 6x^4 - 3x^2 + 2$.

*Solution.* We begin by finding the critical points, so we differentiate to get $f'(x) = 24x^3 - 6x = 6x(4x^2-1) = 6x(2x-1)(2x+1)$. Setting $f'(x) = 0$ and solving for $x$, we get $x = 0, \pm 1/2$. Now since $f'$ splits into linear factors, and each linear factor can only switch sign once, we can determine

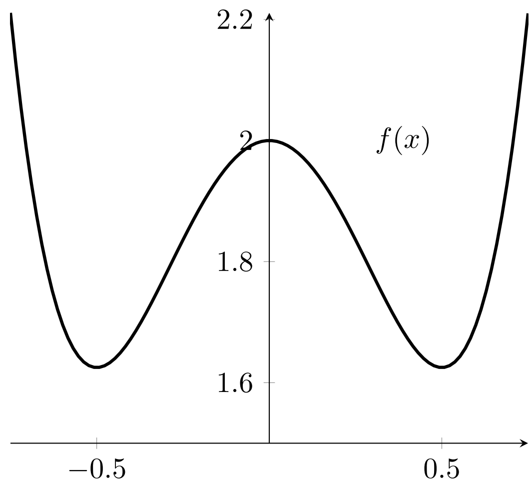



Figure 7.2: The function $f(x) = 6x^4 - 3x^2 + 2$ from Example 7.29.

whether $f$ is increasing or decreasing on each interval by creating the following chart:

| | $x < -\frac{1}{2}$ | $-\frac{1}{2} < x < 0$ | $0 < x < \frac{1}{2}$ | $x > \frac{1}{2}$ |
|---|---|---|---|---|
| $2x+1$ | $-$ | $+$ | $+$ | $+$ |
| $x$ | $-$ | $-$ | $+$ | $+$ |
| $2x-1$ | $-$ | $-$ | $-$ | $+$ |
| $f(x)$ | $-$ | $+$ | $-$ | $+$ |

From our chart we see that $-1/2$ is a min, 0 is a max, and $1/2$ is a min. This is confirmed by the graph of our function. ■

Note however that not all critical points are a max or a min.

**Example 7.30**

What are the critical points of the function $f(x) = x^3$? What are the max/min of $f$?

*Solution.* The derivative of $f$ is $f'(x) = 3x^2$ which has a single zero at $x = 0$ and hence has a single critical point at 0. However, the function $f$ has no local maximum or minimum anywhere, so $x = 0$ is a critical point which is neither a maximum nor a minimum. ■

Theorem 7.27 combined with Example 7.30 imply that while every max/min is a critical point, not all critical points are max/mins. Regardless, critical points provide a powerful tool for finding absolute maxima and minima. As we must also check whether the function achieves its max/min on the boundary, our strategy is as follows:

**Finding Extreme Points**

1. Determine the value of $f$ at the boundary points.
2. Determine the critical points of $f$ by computing the points where $f'(x) = 0$.
3. Evaluate the function $f$ at its critical points.
4. The absolute max and min will be the largest and smallest values from steps (1) and (3).

**Example 7.31**

Determine the global maximum and minimum of the function $f(x) = xe^{-x^2}$ on the interval $[0, 1]$.

*Solution.* Following our algorithm above, we first evaluate $f$ on its endpoints to find that

$$f(0) = 0, \qquad f(1) = \frac{1}{e}.$$

Next we determine the critical points. The derivative of $f$ is given by $f'(x) = e^{-x^2}(1 - 2x^2)$. The component $e^{-x^2}$ is never zero, so the zeroes of $f'$ will occur precisely when $1 - 2x^2 = 0$ which corresponds to $x = \pm 1/\sqrt{2}$. However, notice that $-1/\sqrt{2}$ is not in the interval $[0, 1]$ so we throw it away and just plug $x = 1/\sqrt{2}$ into $f$ to get

$$f\left(\frac{1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}}e^{-\frac{1}{2}} = \frac{1}{\sqrt{2e}}.$$

A quick comparison tells us that $f(0) < f(1) < f(1/\sqrt{2})$ so that 0 is the global minimum and $1/\sqrt{2}$ is the global maximum. ■

**Example 7.32**

Find the global maximum and minimum of the function $f(x) = 3x(x+2)^{2/3}$ on the interval $[-2, 6]$.

*Solution.* Differentiating $f$ gives

$$f'(x) = 3(x+2)^{2/3} + \frac{2x}{(x+2)^{1/3}} = \frac{5x+6}{(x+2)^{1/3}},$$

which has critical points at $x = -6/5$ and $x = -2$. The point $x = -2$ is already an endpoint, so evaluating we get

$$f(-2) = 0, \quad f(-6/5) = -\frac{18}{5}\left(\frac{4}{5}\right)^{1/3} \approx -3.10, \quad f(6) = 18\sqrt[3]{8} = 36.$$

Hence the minimum occurs in the interior at $x = -6/5$, but the maximum occurs at the $x = 6$ endpoint. ■

There is another way to tell whether the endpoints of of function will be a max and min. In the example above, we computed the derivative $f'(x) = x^2/(1+x^2)$ which is always non-negative, and away from $x = 0$ is actually always positive. This means that the function is always increasing, and a function which is always increasing cannot have any max/min on the interior of its domain. With increasing functions, the minimum must occur at the left-most endpoint, while the maximum must occur at the right-most endpoint.

The algorithm given above is excellent for determining global maxima and minima, but what if we want to determine which critical points are local extreme points?

**Theorem 7.33: Second Derivative Test**

Let $f$ be a function which is twice continuously differentiable in a neighbourhood of a critical point $c$.

1. If $f''(c) > 0$ then $c$ is a local minimum.
2. If $f''(c) < 0$ then $c$ is a local maximum.
3. If $f''(c) = 0$ let $k$ be the smallest positive integer such that $f^{(k)}(c) \neq 0$. If $k$ is odd then $c$ is neither a max nor a min. If $k$ is even and $f^{(k)}(c) > 0$ then $c$ is a local max, while if $f^{(k)}(c) < 0$ then $c$ is a local min.

Note the similarities between this and Proposition 7.21. Indeed, the Second Derivative Test simply states that if your function is concave up at a critical point, it must have been a minimum; while if your function is concave down, your critical point must have been a maximum.

**Example 7.34**

Classify all maxima/minima of the function $f(x) = 2x^3 - 3x^2 - 12x + 10$.

*Solution.* We determine the critical points via the derivative, which we compute to be $f'(x) = 6x^2 - 6x - 12 = 6(x^2 - x - 2) = 6(x-2)(x+1) = 0$. This has roots at $x = 2$ and $x = -1$. In order to determine whether these are local maxima or minima, we use the second derivative test. We compute $f''(x) = 12x - 6 = 6(2x-1)$, which evaluated at the critical points gives

$$f''(2) = 18 > 0 \quad \text{and} \quad f''(-1) = -18 < 0,$$

which implies that $x = 2$ is a minimum while $x = -1$ is a maximum. ∎

**Example 7.35**

Check that the critical point we found in Example 7.32 is a minimum using the second derivative test.

*Solution.* Recalling example 7.32, the function $f$ and its derivative $f'$ are given by

$$f(x) = 3x(x+2)^{2/3} \quad \text{and} \quad f'(x) = \frac{5x+6}{(x+2)^{1/3}}.$$

Differentiating $f'$ again gives

$$f''(x) = \frac{2}{3}\frac{5x+12}{(x+2)^{4/3}}.$$

Evaluating $f''(-6/5) \approx 5.39$ shows that $x = -6/5$ is a minimum, as expected. ■

### 7.4.1 Optimization

Here we consider constrained optimization, where our goal is to maximize/minimize a function subject to constraints. There are entire fields of mathematics and engineering dedicated entirely to the study of optimization. The principal difference between an optimization problem and a max/min problem is that optimization problems have *constraints*; that is, we are asked to find the maximize/minimize a quantity, but only so long as some other condition is satisfied. The addition of constraints typically makes the problem more difficult (or if you're a mathematician, much more interesting).

Keeping with the mathematical philosophy that one should reduce new problems to problems which we have already solved, our goal will be to take constrained optimization problems and turn them into simple max/min problems. Typically, such questions will be posed as word-problems amounting to the following:

"Maximize $f(x_1, \ldots, x_n)$ subject to the constraint that $g(x_1, \ldots, x_n) = c$."

Written abstractly, it's not clear how we should proceed. The trick is to use the constraint to rewrite $f(x_1, \ldots, x_n)$ as a function of a single variable, to which we can then apply our usual techniques. Let's look at an example.

**Example 7.36**

An exiled queen and her entourage washed up on the coast of North Africa, and pleaded with a local king to be given a small plot of land that she might catch her breath and rebuild. The king agreed, and they settled on however so much land could be encompassed by a piece of oxhide. The clever queen cut the oxhide into many small strips, so that they stretched a distance of 800 metres. She then also settled beside a straight river so that the river formed one of the barriers of her new kingdom. Given that her primitive people can only lay the oxhide strips in straight lines, find the dimensions of the corresponding rectangle that maximizes the area of her new queendom.[a]

[a] This is a modification of the problem of Queen Dido and the founding of Carthage. Dido's problem is also called an *isoperimetric inequality*. I have made modifications here so that we can solve it using the tools we have developed, but a proper treatment requires the *calculus of variations*.

*Solution.* Stripping away the extra information, we are told to build three sides of a rectangle which maximizes the area of the rectangle, ensuring that the perimeter is 800 metres. Let $y$ be the length of the side opposite the rivers, and $x$ be the length of the two sides perpendicular to the river. Our perimeter is thus $2x + y = 800$ while our area is $A = xy$; that is,

"Maximize the function $f(x, y) = xy$ subject to the constraint $2x + y = 800$."

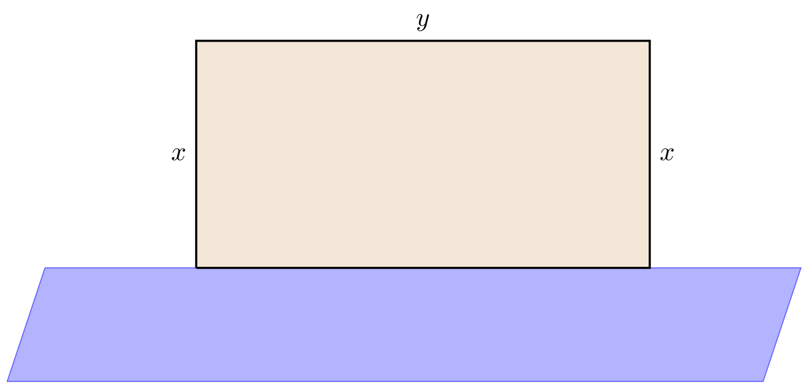



Figure 7.3: Establishing Queen Dido's kingdom, using the river as one of the edges.

As mentioned earlier, the constraint can be manipulated so that $f(x, y) = xy$ becomes an equation of only one variable. Indeed, $2x + y = 800$ implies that $y = 800 - 2x$. Substituting this into $f(x, y)$ we get

$$\hat{f}(x) = f(x, y) = xy = x(800 - 2x) = 800x - 2x^2.$$

This new function already has the information about the constraint encoded into it, so we now content ourselves to simply find the maximum of $\hat{f}(x)$. Finding the critical points we get $\hat{f}'(x) = 800 - 4x = 0$ implies that $x = 200$. We can now use our constraint $2x + y = 800$ to see that $y = 400$. To see that this is indeed a maximum, we note that $\hat{f}''(x) = -4 < 0$, so the second derivative test verifies maximality. ■

The solution is somewhat counter-intuitive. We found that the rectangle we should build is twice as long as it is wide, but if we had not built along the side of the river then solving the optimization problem would have revealed that the optimal area is given by a square (try this on your own, to see that $x = y = 200$). The fact that we built along the river means that our constraint equation changed, and that change propagated to the solution.

Sometimes the constraints can be avoided by being clever, as the following example demonstrates.

**Example 7.37**

Consider a segment of string of length 20 centimetres. If we cut this string into two segments and from from each segment a square and an equilateral triangle, find the cuts which will both maximize and minimize the sum of the areas of each shape.

*Solution.* Let the side length of our square be given by $x$ and the side length of the triangle by $y$. Our constraint implies that the sum of the perimeters of these shapes must be 20 centimetres, so $4x + 3y = 20$. Now one can calculate that the area of an equilateral triangle of side-length $y$ is $A_\Delta = \frac{\sqrt{3}}{4}y^2$ and the area of the square is obviously $A_\Box = x^2$. Hence our problem is

"Maximize/minimize $f(x, y) = A_\Delta + A_\Box = x^2 + \frac{\sqrt{3}}{4}y^2$, subject to the constraint $4x + 3y = 20$."

As before, our constraint allows us to rewrite $f(x, y)$ as a function of just one variable. The equation $4x + 3y = 20$ implies that $y = (20 - 4x)/3$, and substituting this into $f(x, y)$ gives

$$f(x) = x^2 + \frac{\sqrt{3}}{4}\left(\frac{20-4x}{3}\right)^2 = x^2 + \frac{1}{12\sqrt{3}}\left(400 - 160x + 16x^2\right).$$

Differentiating to find the critical points, we get

$$0 = f'(x) = 2x + \frac{1}{12\sqrt{3}}(-160 + 32x)$$

which we can solve to get $x = 20/(3\sqrt{3} + 4)$. This implies that

$$y = \frac{20-4x}{3} = \frac{20\sqrt{3}}{3\sqrt{3}+4}.$$

Hence the side lengths of the square and triangle are

$$\frac{80}{3\sqrt{3}+4} \quad \text{and} \quad \frac{60}{3\sqrt{3}+4} \quad \text{respectively.} \tag{7.6}$$

You can use a calculator to see that this corresponds to an approximate area of 10.87 cm$^2$. This is also a local minimum since it is easy to see that the second derivative of this function is always positive. By our previous treatment of max/min, we know that we must also check the endpoints for a solution. If $x = 0$ then $y = 20/3$ gives an area of 19.25 while if $y = 0$ then $x = 5$ gives an area of 25. Hence our area is maximized when we use the string to only make the square, and minimized with the side lengths given in (7.6). ■

## 7.5 Curve Sketching

The grand-total of all the tools hitherto developed give us the ability to analyze functions and determine the behaviour of their graphs. One of the important applications of this is that while we may implement Computer Algebra Systems to help us analyze functions, it is still essential for the operator to understand the fundamentals in order to find things that a computer would otherwise miss. A simple but important example is as follows: Consider the function $f(x) = \frac{1}{300}(x^4 - 2x^2 + 1)$. If we were to graph this using software, we might get the graph in Figure 7.4: (Left).

Now let us assume that this figure describes the potential energy of a system. Any state will try to minimize its potential energy, and so it is tempting to assume that the point at $x = 0$ describes a global (stable) minimum and so would be an excellent place to initialize a state. However, the use of calculus actually shows that the point $x = 0$ is a local maximum and hence is an unstable equilibrium. If implemented as an engineering solution, this could quickly lead to disaster.

Our goal is thus to combine all of our information into a system which allows us to analyze the qualitative behaviour of a function without knowing the nitty-gritty details of its exact value at every point. There are approximately seven pieces of information that we need to compute to ascertain the general behaviour.

1. Domain (and range if possible),

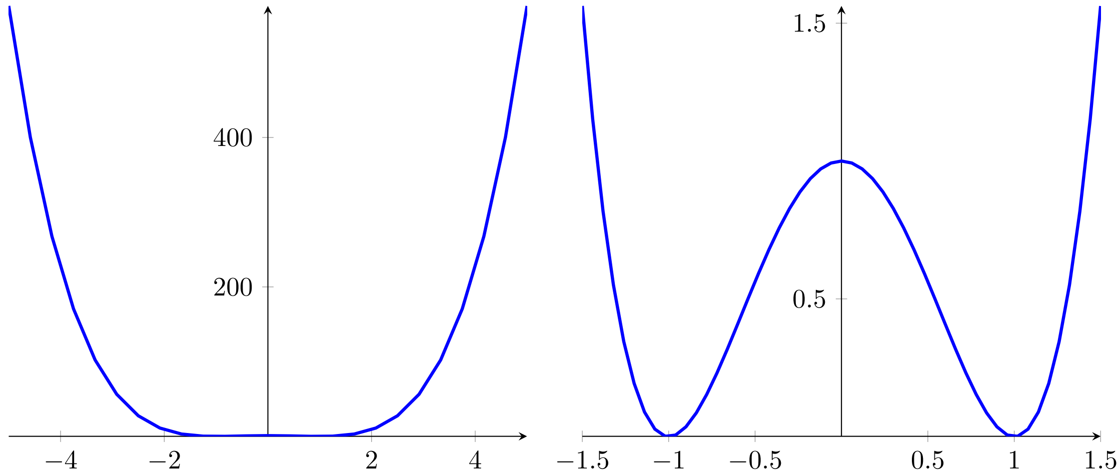



Figure 7.4: The graph of a misleading function. **Left:** The graph of our function as given by a computer. **Right:** Zooming in, we see the presence of a local maximum at $x = 0$.

2. Intercepts ($x$- and $y$-),
3. Symmetry (even/odd/none),
4. Asymptotes (horizontal/vertical/oblique),
5. Increasing (and obviously decreasing),
6. Maxima (and minima),
7. Concavity and inflection points.

Of course, there is one piece of information above that I have not included; namely, what is an oblique (aka slant) asymptote?

**Definition 7.38**

Let $f(x)$ and $g(x)$ be continuous functions. We say that $f(x)$ *behaves like* $g(x)$ *asymptotically* if

$$\lim_{x\to\pm\infty} [f(x) - g(x)] = 0.$$

We say that $f(x)$ has an *oblique asymptote* if $f(x)$ behaves asymptotically like $g(x) = mx+b$ for some $m \neq 0$.

**Example 7.39**

Compute any asymptotics of the function $f(x) = \dfrac{2x^3 - x^2 + 2x}{x^2+1}$.

*Solution.* The easiest way to proceed is to try to write $f$ as an improper rational function. Performing long division, we see that

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

The idea is that in the limit as $x \to \infty$, the $1/(x^2+1)$ term will die off and contribute very little to the behaviour of $f(x)$, so that $f(x)$ looks like the function $2x-1$. To see that this satisfies the definition above, set $g(x) = 2x-1$ so that

$$\lim_{x\to\infty} [f(x) - g(x)] = \lim_{x\to\infty} \left[2x - 1 + \frac{1}{x^2+1} - (2x-1)\right] = \lim_{x\to\infty} \frac{1}{x^2+1} = 0,$$

which is precisely what we wanted to show. Similarly, the limit as $x \to -\infty$ shows that $f$ is also asymptotically like $g$ in that limit as well. Thus $2x-1$ is an oblique asymptote for $f$ at both $\pm\infty$. ∎

Now that we know how to compute all terms involved in this computation, we shall proceed with some examples.

**Example 7.40**

Plot the function $f(x) = \dfrac{x^3}{(x+1)^2}$.

*Solution.* **Domain:** The only point which could possibly give us trouble is $x = -1$. Hence our domain is simply $\mathbb{R} \setminus \{-1\}$.

**Intercepts:** The $y$-intercept occurs when $x = 0$, so namely $f(0) = 0$. Similarly the $x$-intercept comes when $y = 0$, for which we see that

$$\frac{x^3}{(x+1)^2} = 0 \qquad \Leftrightarrow \qquad x = 0.$$

Thus the $x$- and $y$-intercepts both occur at the origin.

**Symmetry:** There is no symmetry involved: Since the functions are polynomial they have no periodicity. You can check that $f(-x)$ has no relation to $f(x)$, so that the function is neither even nor odd.

**Asymptotes:** We begin with the horizontal asymptotes. It is easy to see that since the numerator dominates the denominator, the limit will go to infinity (check this by dividing top and bottom by $1/x^3$). Further, since the denominator is always positive, the sign is determined entirely by the $x^3$ factor, so

$$\lim_{x\to\infty} \frac{x^3}{(x+1)^2} = \infty, \qquad \lim_{x\to-\infty} \frac{x^3}{(x+1)^2} = -\infty.$$

We conclude there are no horizontal asymptotes.

The only candidate for a vertical asymptote occurs at $x = -1$. Again the denominator $(x+1)^2$ is always positive, so the sign of the "infinity" is entirely determined by the behaviour of $x^3$ around

$x = -1$, which is negative. It is then clear that

$$\lim_{x \to -1^\pm} \frac{x^3}{(x+1)^2} = -\infty.$$

Finally, we want to check for oblique asymptotes. Using long polynomial division we may easily find that

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

so we claim that $y = x - 2$ is an oblique asymptote. Indeed, notice that

$$\begin{aligned}
\lim_{x \to \pm\infty} [f(x) - (x-2)] &= \lim_{x \to \pm\infty} \left[ \left( (x-2) - \frac{3x+2}{x^2+2x+1} \right) - (x-2) \right] \\
&= - \lim_{x \to \pm\infty} \frac{3x+2}{x^2+2x+1} \\
&= - \lim_{x \to \pm\infty} \frac{3/x + 2/x^2}{1 + 2/x + 1/x^2} && \text{multiply and divide by } 1/x^2. \\
&= 0.
\end{aligned}$$

**First Derivative:** Computing the first derivative can be a chore, but we find that

$$\begin{aligned}
f'(x) &= \frac{3x^2(x+1)^2 - 2(x+1)x^3}{(x+1)^4} = \frac{3x^4 + 6x^3 + 3x^2 - 2x^4 - 2x^3}{(x+1)^4} \\
&= \frac{x^2(x+3)(x+1)}{(x+1)^4} = \frac{x^2(x+3)}{(x+1)^3}
\end{aligned}$$

so that the critical points correspond to $x = -1$, $x = 0$ and $x = -3$. The $y$-values for these points will be useful when we plot, so we substitute to find that $f(0) = 0$ and $f(-3) = 27/16$. To determine where the function is increasing and decreasing, we consider the following table:

| | $x < -3$ | $-3 < x < -1$ | $-1 < x < 0$ | $0 < x$ |
|---|---|---|---|---|
| $x+3$ | $-$ | $+$ | $+$ | $+$ |
| $(x+1)^3$ | $-$ | $-$ | $+$ | $+$ |
| $x^2$ | $+$ | $+$ | $+$ | $+$ |
| $f(x)$ | $+$ | $-$ | $+$ | $+$ |

**Second Derivative:** The second derivative is a little messy, but simplifies if done correctly.

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x} \frac{x^2(x+3)}{(x+1)^3} &= \frac{(3x^2+6x)(x+1)^3 - 3(x+1)^2(x^3+3x^2)}{(x+1)^6} \\
&= \frac{3x^3 + 3x^2 + 6x^2 + 6x - 3x^3 - 9x^2}{(x+1)^4} \\
&= \frac{6x}{(x+1)^4}
\end{aligned}$$

so there is an inflection point at $(0,0)$ (telling us that one of the critical points is an inflection point. Furthermore, $f''(-3) = -9/4 < 0$ so the point $(3, 27/16)$ is a max. Since the denominator is a quartic it is always positive, and we can see that concavity is entirely determined by the numerator $6x$. Hence $f(x)$ is concave up when $f''(x) > 0$, corresponding to $x > 0$; and $f(x)$ is concave down when $f''(x) < 0$, corresponding to $x < 0$.

**Plotting:** Putting all of this information together, you will get Figure 7.5.

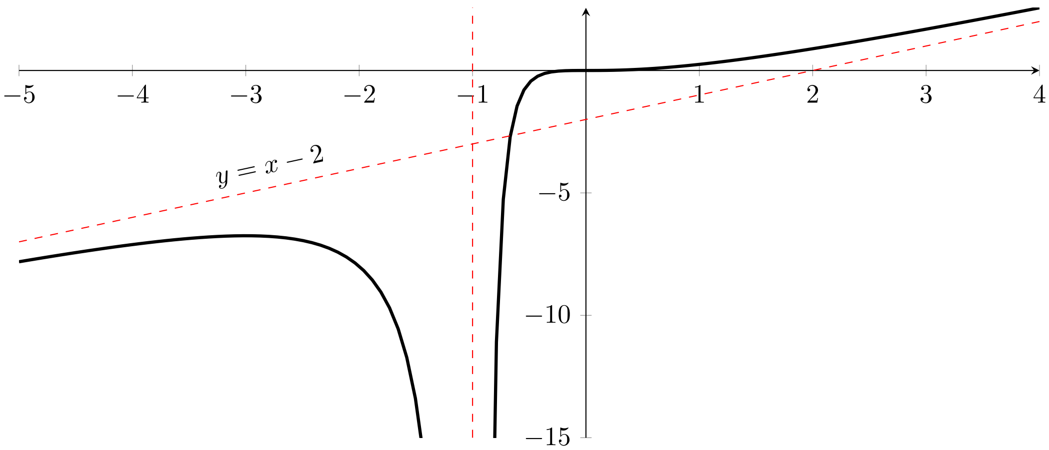



Figure 7.5: A plot of the curve $f(x) = \dfrac{x^3}{(1+x)^2}$.

■

**Example 7.41**

Plot the function $f(x) = \dfrac{x^3}{1-x^2}$.

*Solution.* Following our program we set to work.

**Domain:** By this point, this should not be too hard to see. In particular, our function will not be defined whenever the denominator is zero. This happens at the points $x = \pm 1$ and so our domain is $\mathbb{R} \setminus \{\pm 1\}$.

**Intercepts:** The $y$-intercept occurs when $x = 0$, so namely $f(0) = 0$. Similarly the $x$-intercept comes when $y = 0$, for which we see that

$$\frac{x^3}{1-x^2} = 0 \qquad \Leftrightarrow \qquad x = 0.$$

Thus the $x$- and $y$-intercepts both occur at the origin.

**Symmetry:** Since we are dealing with polynomials, there is no obvious periodicity to worry about. It's not too hard to see that this is actually an odd function, since

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

**Asymptotes:** The vertical asymptotes will clearly occur at $x = \pm 1$. Typically, one would calculate the limits

$$\lim_{x\to 1^\pm} \frac{x^3}{1-x^2}, \quad \lim_{x\to -1^\pm} \frac{x^3}{1-x^2}$$

but this is laborious and is redundant once we have information on the first derivative. For the interested student who would like to see how to do this all the same, we have the following table

| | $x^3$ | $1-x^2$ | $x^3/(1-x^2)$ |
|---|---|---|---|
| $x \to 1^+$ | $+$ | $-$ | $-$ |
| $x \to 1^-$ | $+$ | $+$ | $+$ |
| $x \to -1^+$ | $-$ | $+$ | $-$ |
| $x \to -1^-$ | $-$ | $-$ | $+$ |

so that

$$\lim_{x\to 1^-} \frac{x^3}{1-x^2} = \lim_{x\to -1^-} \frac{x^3}{1-x^2} = \infty \quad \text{and} \quad \lim_{x\to 1^+} \frac{x^3}{1-x^2} = \lim_{x\to -1^+} \frac{x^3}{1-x^2} = \infty$$

Because the degree of the numerator is strictly greater than the degree of the denominator, there are no horizontal asymptotes:

$$\lim_{x\to\pm\infty} \frac{x^3}{1-x^2} = \mp\infty.$$

Finally, we want to check for oblique asymptotes. Using long polynomial division we may easily find that

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

so we claim that $y = -x$ is an oblique asymptote. Indeed, notice that

$$\begin{aligned}\lim_{x\to\pm\infty} [f(x) - (-x)] &= \lim_{x\to\pm\infty} \left[\frac{x^3}{1-x^2} + x\right] = \lim_{x\to\pm\infty} \frac{x}{1-x^2} \\ &= \lim_{x\to\pm\infty} \frac{1/x}{1/x^2 - 1} = 0.\end{aligned}$$

**First Derivative:** This step allows us to determine where the function is increasing, decreasing, the critical points, and when combined with the second derivative, maxima and minima. The first derivative is computed to be

$$\frac{\mathrm{d}}{\mathrm{d}x} \frac{x^3}{1-x^2} = \frac{(3x^2)(1-x^2) - (-2x)(x^3)}{(1-x^2)^2} = \frac{x^2(3-x^2)}{(1-x^2)^2}.$$

We may simply read off the critical points as $\left\{\pm 1, \pm\sqrt{3}, 0\right\}$ with potential extrema at $\left\{0, \pm\sqrt{3}\right\}$. Setting up a quick table for increasing and decreasing we have

| | $x < -\sqrt{3}$ | $-\sqrt{3} < x < -1$ | $-1 < x < 0$ | $0 < x < -1$ | $1 < x < \sqrt{3}$ | $x > \sqrt{3}$ |
|---|---|---|---|---|---|---|
| $f(x)$ | $-$ | $+$ | $+$ | $+$ | $+$ | $-$ |

This table is easy to deduce once we realize that the $\frac{x^2}{(1-x^2)^2}$ portion is always positive, so the sign of $f'(x)$ is entirely determined by the sign of $3-x^2$, which is negative whenever $|x| > \sqrt{3}$. It will likely be useful to know the function values corresponding to our critical points. We already know that $f(0) = 0$ and we find that

$$f(\pm\sqrt{3}) = \frac{\pm 3\sqrt{3}}{-2} = \frac{\mp 3\sqrt{3}}{2}.$$

**Second Derivative:** The second derivative is a little messy, but simplifies if done correctly.

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{x^2(3-x^2)}{(1-x^2)^2} = \frac{(6x-4x^3)(1-x)^2 - 2(1-x^2)(-2x)(3x^3-x^4)}{(1-x^2)^4} = \frac{2x(x^2+3)}{(1-x^2)^3}.$$

The inflection points will occur when $f''(x) = 0$ or does not exist, which we can again read off as being $\{0, \pm 1\}$. We form a table to check for concavity and find

| | $x < -1$ | $-1 < x < 0$ | $0 < x < 1$ | $x > 1$ |
|---|---|---|---|---|
| $f(x)$ | $+$ | $-$ | $+$ | $-$ |

Finally, recalling that we have extrema candidate at $\pm 3$ we check to find that

$$f(\pm\sqrt{3}) = \frac{\pm 2\sqrt{3}(3+3)}{(1-3)^2} = \pm\frac{12\sqrt{3}}{-8} = \mp\frac{3\sqrt{3}}{2}.$$

Thus $\left(\sqrt{3}, -\frac{3\sqrt{3}}{2}\right)$ is a local maximum and $\left(-\sqrt{3}, \frac{3\sqrt{3}}{2}\right)$ is a local minimum. Since $f''(0) = 0$ we cannot infer any information about this critical point. If we continue to take derivatives, we will find that $f^{(3)}(0) = 6$ and so by the generalized second derivative test, 0 is an inflection point.

**Plotting:** Putting all of this information together, the student should get the following plot:

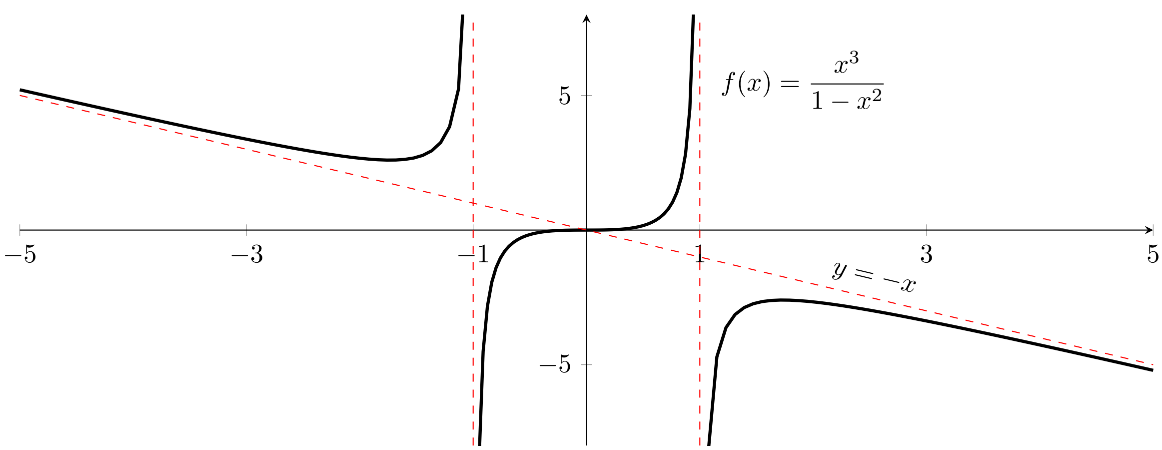



Figure 7.6: A plot of the curve $f(x) = \dfrac{x^3}{1-x^2}$.

■

## 7.6 Approximating a Function

The simplest non-trivial approximation to a function is a first order approximation, sometimes referred to as the *linear approximation to $f$ at $a$*. The derivative $f'(a)$ represents the slope of the tangent line to $f$ at $a$, which should be a reasonable approximation to $f$ near $a$. This line has the form

$$p_{1,a}(x) = f(a) + f'(a)(x - a). \tag{7.7}$$

Geometrically, as long as we stay near the point $a$ the straight line should do a decent job of approximating the function, such as in Figure 7.7

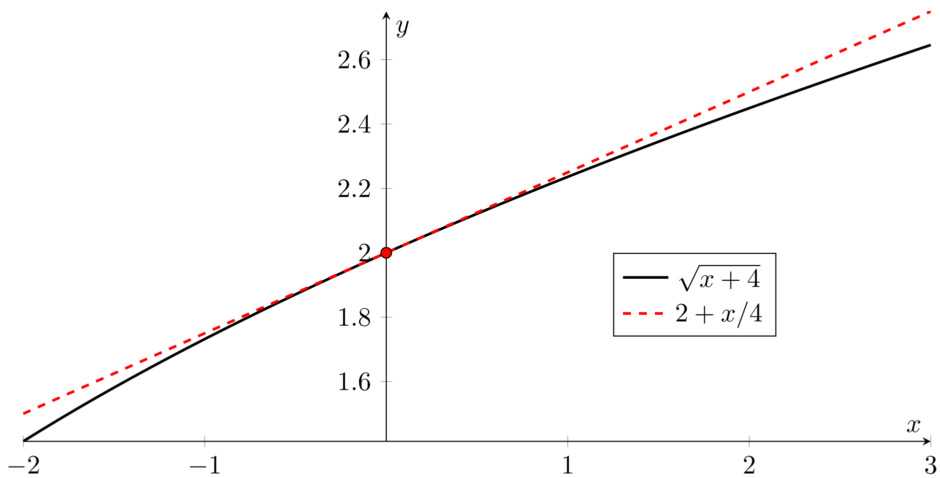



Figure 7.7: The linear approximation to the function $f(x) = \sqrt{x+4}$ at $x = 0$ is $p_{1,0}(x) = 2 + x/4$.

**Example 7.42**

Use linear approximation to find estimates for the numbers $\sqrt{4.1}$, $\sqrt{4.2}$, $\sqrt{5}$, $\sqrt{6}$.

*Solution.* There are several possible functions which we could use for our approximation, though the two simplest choices are

$$f(x) = \sqrt{x+4} \text{ at } x = 0 \quad \text{or} \quad g(x) = \sqrt{x} \text{ at } x = 4.$$

Let's proceed using $f$, but the answers will be identical were we to use $g$ instead. Since $f'(x) = [2\sqrt{x+4}]^{-1}$ we have that $f'(0) = 1/4$ and the linear approximation is thus

$$p_{1,0}(x) = f(0) + f'(0)(x - 0) = 2 + \frac{x}{4}.$$

To approximate $\sqrt{4.1}$, we know that $f(0.1) = \sqrt{4 + 0.1} = \sqrt{4.1}$ and that

$$f(0.1) \approx p_{1,0}(0.1) = 2 + \frac{1}{4}(0.1) = 2.025.$$

The actual value of $\sqrt{4.1}$ is 2.0248... which amounts to an error of $1.54 \times 10^{-4}$. That is pretty good! The remaining values are given in the following chart:

| Value | $p_{1,0}(x)$ | Actual | Error $= \lvert f(x) - p_{1,0}(x) \rvert$ |
|---|---|---|---|
| $\sqrt{4.1}$ | 2.025 | 2.0248 | $1.54 \times 10^{-4}$ |
| $\sqrt{4.2}$ | 2.05 | 2.0494 | $6.10 \times 10^{-4}$ |
| $\sqrt{5}$ | 2.25 | 2.2361 | $1.39 \times 10^{-2}$ |
| $\sqrt{6}$ | 2.5 | 2.4495 | $5.05 \times 10^{-2}$ |

Note that as we get further away from the approximation point $x = 0$, our errors become worse. ■

**Example 7.43**

Use a first order Taylor polynomial to approximate the value of $\ln(1.5)$.

*Solution.* Let $f(x) = \ln(1 + x)$ and take an approximation at $x = 0$. The derivative is $f'(x) = 1/(1 + x)$, so that $f'(0) = 1$, implying that our linear approximation is $p_{0,1}(x) = x$. This gives us the approximation

$$\ln(1.5) = f(0.5) \approx p_{0,1}(0.5) = 0.5.$$

This doesn't feel like a very satisfying answer, and when we compare it to the true value of $\ln(1.5) = 0.4055...$, we are even less satisfied. For practical purposes we'll need to use a higher order approximations. We'll use a cubic approximation in Example 7.44, which will do a much better job. ■

### 7.6.1 Quadratic and Higher

By using more terms in the polynomial we get even better approximations. The formula for the $n$th order approximation to $f$ at $x = a$ is

$$\begin{aligned} p_{n,a}(x) &= \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1} + \cdots + \frac{f''(a)}{2}(x-a)^2 + f'(a)(x-a) + f(a) \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k. \end{aligned}$$

For example, with the function $f(x) = e^x$, we get the following approximations at 0:

| Order | Polynomial |
|---|---|
| 1 | $p_{1,0}(x) = 1 + x$ |
| 2 | $p_{2,0}(x) = 1 + x + \frac{1}{2}x^2$ |
| 3 | $p_{3,0}(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3$ |
| 4 | $p_{4,0}(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4$ |

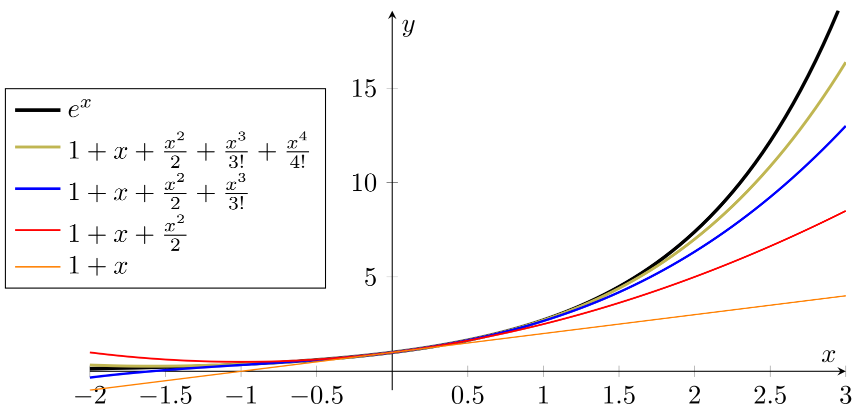



Figure 7.8: Several polynomial approximations to the function $f(x) = e^x$. Notice that each successive polynomial yields a better approximation.

**Example 7.44**

Use a cubic approximation to estimate the value of $\ln(1.5)$.

*Solution.* Set $f(x) = \ln(1 + x)$ and take the approximation at $x = 0$. Here we get

$$\begin{aligned} f(0) &= \ln(1) = 0 \\ f'(0) &= \left[\frac{1}{1+x}\right]_{x=0} = 1 \\ f''(0) &= \left[-\frac{1}{(1+x)^2}\right]_{x=0} = -1 \\ f^{(3)}(0) &= \left[\frac{2}{(1+x)^3}\right]_{x=0} = 2 \end{aligned}$$

so the cubic approximation to $f$ at $x = 0$ is

$$p_{3,0}(x) = f(0) + f'(0)x + \frac{f''(0)}{2} + \frac{f^{(3)}(0)}{6} = x - \frac{1}{2}x^2 + \frac{1}{3}x^3.$$

To estimate $\ln(1.5)$ we get

$$\ln(1.5) = f(0.5) = (0.5) - \frac{1}{2}(0.5)^2 + \frac{1}{3}(0.5)^3 = 0.41\bar{6}.$$

This estimate has an absolute error of approximation $1.1 \times 10^{-2}$, which is much better than what we found in Example 7.43. ■

**Example 7.45**

Use a fourth order approximation on $f(x) = e^x$ to find an estimate for $e$.

*Solution.* Set $f(x) = e^x$ so that $e = f(1)$. We'll approximate $e$ by finding a fourth-order Taylor polynomial for $f$ at $x = 0$, and evaluating this at $x = 1$. The derivatives of $f$ are simple to compute, giving $f^{(k)}(x) = e^x$. Hence $f^{(k)}(0) = 1$ for every $k$, and the fourth order Taylor polynomial is

$$\begin{aligned} p_{4,0}(x) &= f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \frac{f^{(4)}(0)}{4!}x^4 \\ &= 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24}. \end{aligned}$$

Our estimate for $e$ is thus

$$e = f(1) \approx p_{4,0}(1) = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} = 2.708\bar{3},$$

which has an error of $9.95 \times 10^{-3}$. ■

## 7.7 Exercises

7-1. For each given equation, find $\dfrac{dy}{dx}$.

(a) $x\ln(xy) + x + y = 0$

(b) $x^2y^2 + xy = \dfrac{1}{xy}$

(c) $xe^y + ye^x = 0$

(d) $\dfrac{x+y}{xy} = 1$

7-2. For each equation given in Question 1, find $\dfrac{dx}{dy}$.

7-3. Find $\dfrac{d^2y}{dx^2}$ if $x^4 + y^4 = 16$

7-4. Use implicit differentiation to find an equation of the tangent line to the curve $x^{\frac{2}{3}} + y^{\frac{2}{3}} = 4$ at the point $(-3\sqrt{3}, 1)$.

7-5. If $xy + e^y = e$, find $\dfrac{d^2y}{dx^2}$ when $x = 0$.

7-6. Find the points on the curve, $2(x^2 + y^2)^2 = 25(x^2 - y^2)$, where the tangent is horizontal.

7-7. Consider the equation $x^2y^2 - 2xy = 8$.

(a) If $x = 1$, determine the corresponding value(s) for $y$ that satisfy the above equation.

(b) For each point $(x_0, y_0)$ you found in part (a), determine $\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{(x_0,y_0)}$.

(c) For each point $(x_0, y_0)$ you found in part (a), determine $\left.\frac{\mathrm{d}x}{\mathrm{d}y}\right|_{(x_0,y_0)}$

7-8. Consider the equation $\ln(x + y^2) + x^2e^y = 1$.

(a) If $x = 0$, determine the corresponding value(s) for $y$ that satisfy the above equation.

(b) For each point $(x_0, y_0)$ you found in part (a), determine $\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{(x_0,y_0)}$.

(c) For each point $(x_0, y_0)$ you found in part (a), determine $\left.\frac{\mathrm{d}x}{\mathrm{d}y}\right|_{(x_0,y_0)}$

7-9. Consider the equation $x\ln(x + y^2) + xy - y = \ln(5)$.

(a) Find the value(s) of $y$ that correspond to $x = 1$.

(b) Determine $\frac{\mathrm{d}x}{\mathrm{d}y}$ when $x = 1$. Remember to plug in the values of $y$ you found in part (a). Thus if you found three values of $y$ in part (a), you should have three answers.

(c) What is the value of $\frac{\mathrm{d}y}{\mathrm{d}x}$? *Hint:* Use your answer from part (b), do not recompute the derivative.

7-10. Consider the equation $xy + x^2 + y = 3$. Find $\frac{\mathrm{d}x}{\mathrm{d}y}$ when $y = 1$. Provide all possible solutions if there are more than one.

7-11. Consider the equation

$$e^{xy} - y^2 + x\ln(y) = 0.$$

(a) Find the value(s) of $x$ which corresponds to $y = 1$.

(b) Find $\frac{\mathrm{d}y}{\mathrm{d}x}$ when $y = 1$. Your answer should be a number.

(c) Find $\frac{\mathrm{d}x}{\mathrm{d}y}$ when $y = 1$. Your answer should be a number. *Note:* The roles of $x$ and $y$ been interchanged.

7-12. Consider the equation $y^x = x^y$.

(a) Determine the value of $y$ if $x = 1$.

(b) Compute $\frac{\mathrm{d}y}{\mathrm{d}x}$ when $x = 1$.

(c) Compute $\frac{\mathrm{d}x}{\mathrm{d}y}$ when $x = 1$.

7-13. (a) Suppose $x = f(t)$ is defined by the ipmlicit function $t^2x + x - t = 0$. Determine $\frac{\mathrm{d}x}{\mathrm{d}t}$.

(b) Suppose $t = g(y)$ is defined by the implicit function $t^2y + t^2 + y - 1 = 0$. Find $\frac{\mathrm{d}t}{\mathrm{d}y}$.

(c) Consider the composition $x = f(g(y))$, where $f$ and $g$ are defined as above. Compute $\frac{\mathrm{d}x}{\mathrm{d}y}$ when $(x, y, t) = (\sqrt{3}/4, -1/2, \sqrt{3})$.

7-14. Consider the equation $xy^3 - xy = x - 1$.

(a) By differentiating implicitly, find an expression for $\frac{\mathrm{d}x}{\mathrm{d}y}$.

(b) Write $x$ an explicit function of $y$ by solving the above equation for $x$. Use this to compute $\frac{\mathrm{d}x}{\mathrm{d}y}$.

(c) Show that your answers from part (a) and (b) are the same.

7-15. Suppose that $f(x, y) = 0$ is an equation, which you differentiate implicitly at $(x, y) = (a, b)$, and get $\left.\frac{dy}{dx}\right|_{x=a,y=b} = C$. What is $\left.\frac{dx}{dy}\right|_{x=a,y=b}$?

7-16. Find and classify the critical points of each function.

(a) $f(x) = x^3 - 3x^2 - 45x + 22$

(b) $f(x) = \dfrac{x}{x^2+1}$

(c) $f(x) = xe^{-x^2}$

(d) $f(x) = x \ln |x|$

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

7-17. Find and classify the critical points of each function.

(a) $f(x) = \sqrt[3]{x}(8 - x)$

(b) $f(x) = \dfrac{x^3}{x^2 - 1}$.

7-18. Determine the global maximum and minimum of $f$ on the given interval.

(a) $f(x) = 2x - \sqrt{x-1}$ on $[1, 5]$.

(b) $f(x) = x + \dfrac{1}{x}$ on $[1/3, 2]$.

(c) $f(x) = (2x - x^2)e^x$ on $[-2, 2]$.

(d) $f(x) = x^2 \ln |x|$ on $[-2, 2]$.

7-19. Let $A$ be some constant and consider the function

$$F(x) = \ln(1 + x) - \frac{Ax}{1 + x} \quad \text{for} \quad x > -1.$$

Determine the global minimum of $F$. Note that your answer should depend on $A$.

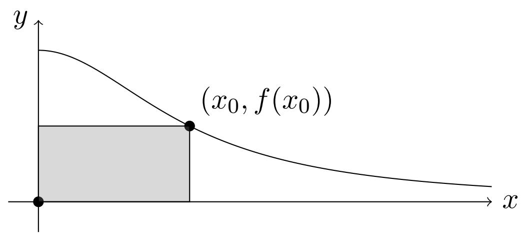
7-20. Suppose $B \in \mathbb{R}$ and consider the function

$$G(x) = Bx^2 - \frac{B + 1}{x}.$$

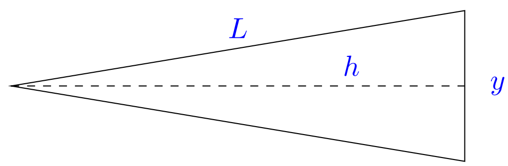
(a) Find the value of $B$ such that $G$ has a critical point at $x = 2$.

(b) Using the value of $B$ you found in part (a), classify the critical point at $x = 2$ as either a maximum, a minimum, or neither.

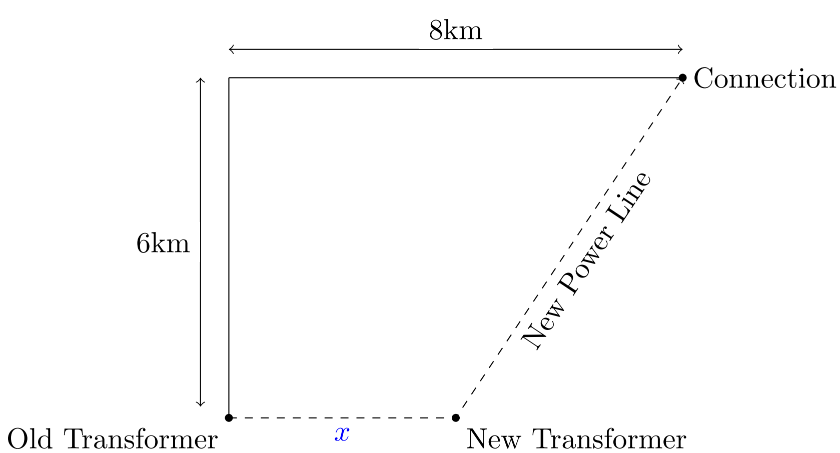
7-21. Consider the function $f(x) = \dfrac{1}{x^2+1}$ with domain $[0, \infty)$. Construct a rectangle with one vertex at the origin $(0, 0)$, and the diagonally opposite vertex on the graph of $f$. Find the value of $x$ which maximizes the area of the rectangle.



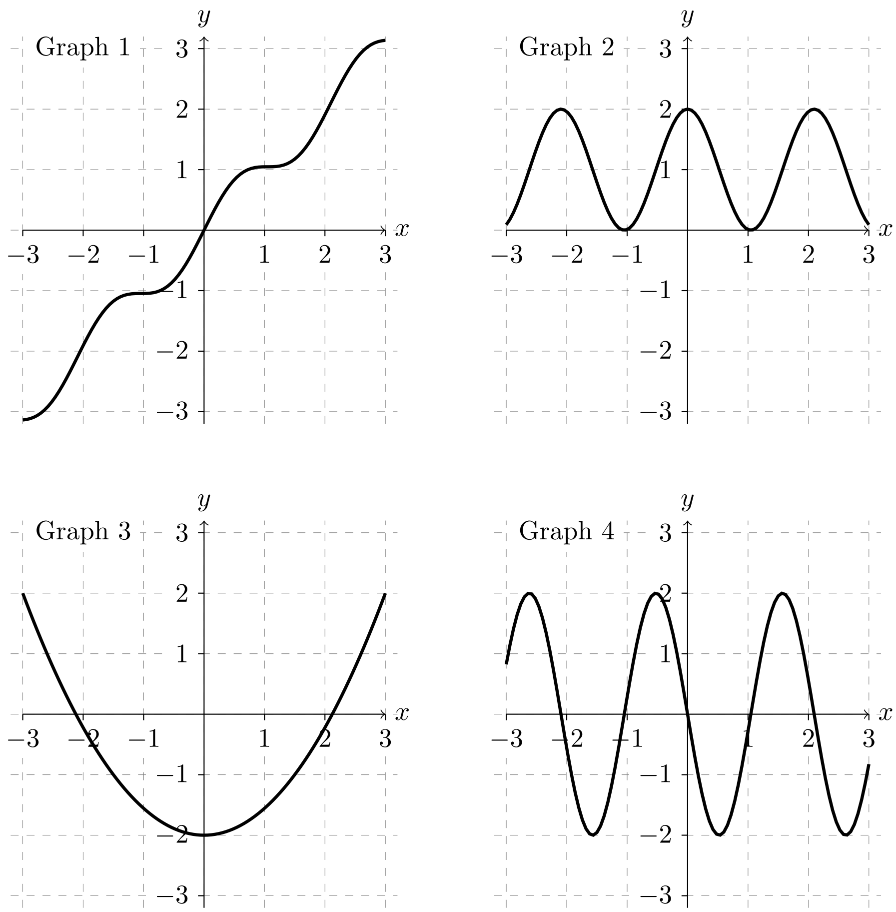

7-22. You are working for a windmill company as their lead designer. The productivity of a windmill is proportionate to the area of its blades, so your goal is to maximize that area. Each windmill you produce must have four identical triangular blades. For engineering reasons, each triangle must be an isosceles triangle, and the total perimeter of each triangular blade cannot be more than 10m. Determine the dimensions of each triangular blade.




7-23. You work for Hydro One, and are faced with the following problem: A transformer you built currently sits in a swamp, needs to be moved, and a new power line established. The transformer sits 6km from the road and must stay 6km from the road. The nearest connection site is 8km away along the road. It will cost you \$3000 for each kilometre you move towards the connection site; that is, it will cost \$3000 for every kilometre between the old transformer and the new transformer. In addition, it costs \$5000 for each kilometre of power line you run from the new transformer to the connnection site. Determine where you should move the transformer to minimize your costs.



7-24. Consider the following graphs




(a) For which graph(s) does the function $f$ satisfy $f'(x) \geq 0$ for all $x > 0$?

(b) For which graph(s) does the function $f$ satisfy $f''(x) \geq 0$ for all $0 \leq x \leq 2$?

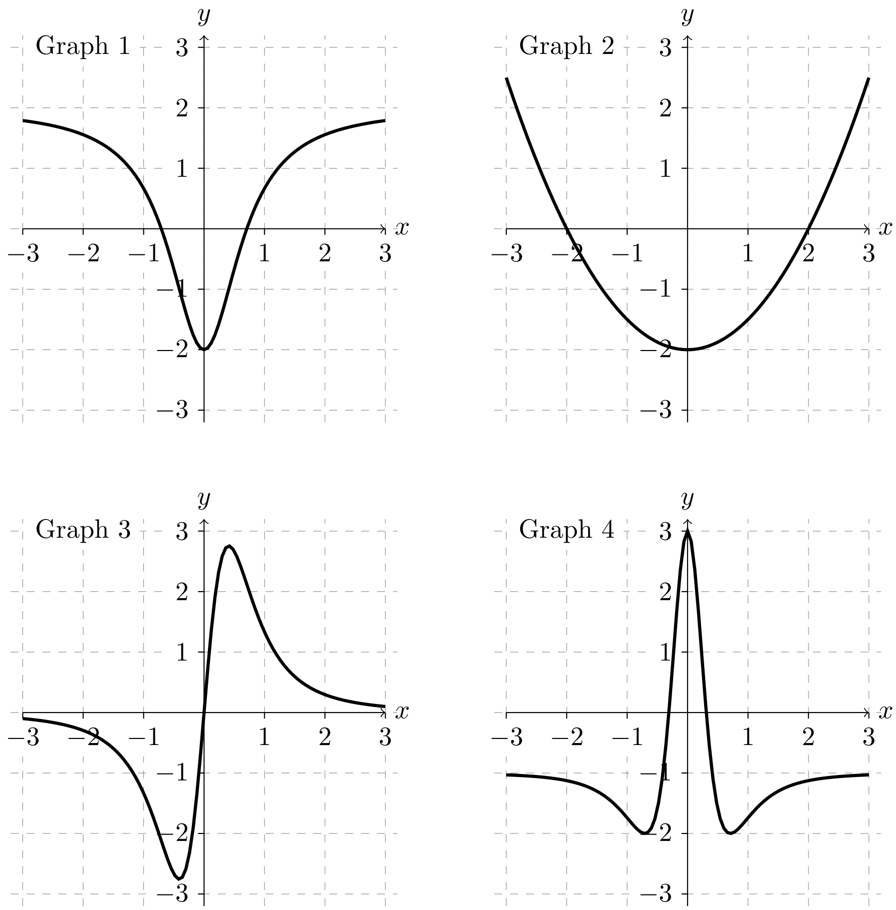
(c) For which graph(s) does $f'$ change sign on $[-1, 1]$?

(d) For which graph(s) does $f''$ change sign on $[-1, 1]$?

(e) Suppose that $f$ is the function given in Graph 1. Which graph corresponds to $f'$?

(f) Suppose that $f$ is the function given in Graph 1. Which graph corresponds to $f''$?

7-25. Consider the following graphs:




(a) Indicate which graph(s) satisfy $f'(x) > 0$ when $x > 0$.

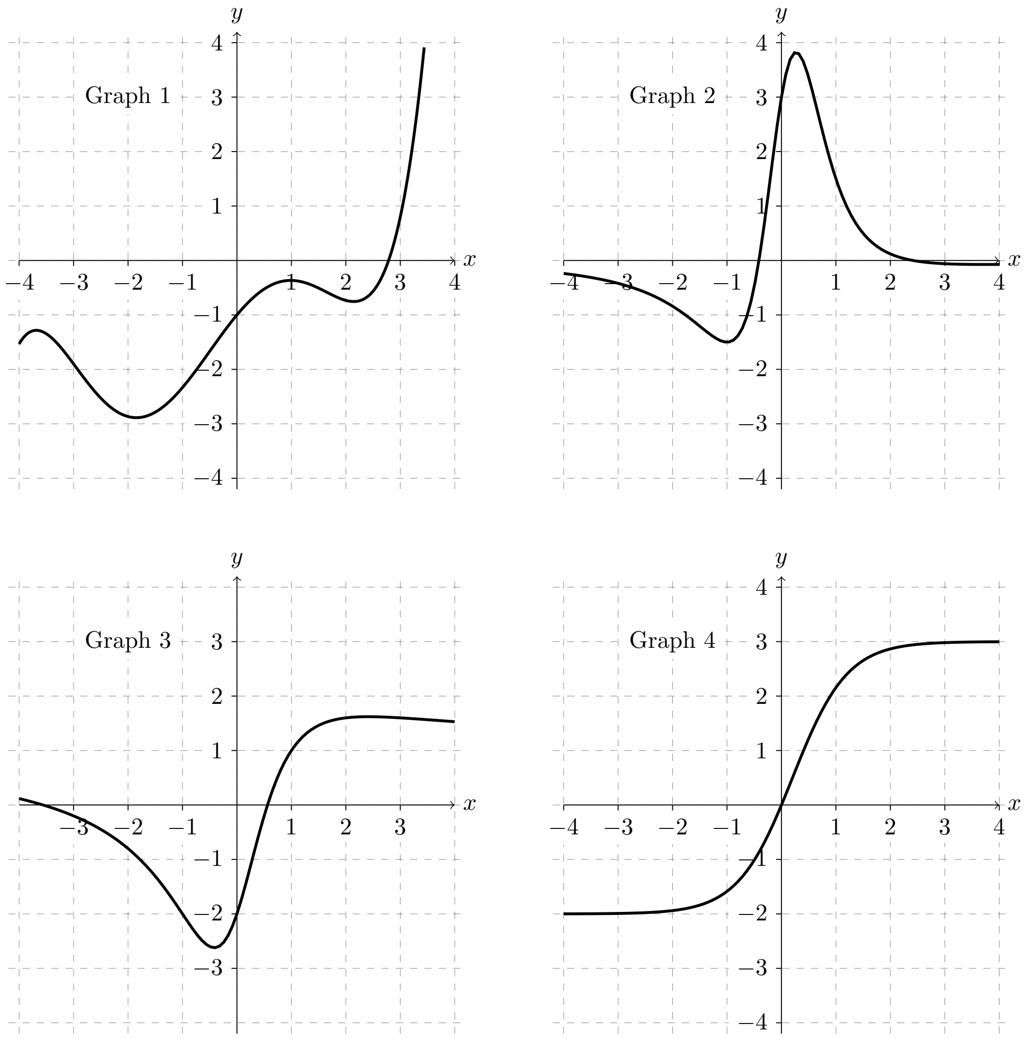
(b) Indicate which graphs change concavity for $x < 0$.

(c) If $f$ is the function illustrated in Graph 1, which graph corresponds to $f'$?

(d) If $f$ is the function illustrated in Graph 1, which graph corresponds to $f''$?

7-26. Let $f$ be the function shown in Graph 4 of Exercise 7-25. Draw the graph of $f'$. To do this, indicate when $f'$ is positive, negative, increasing, decreasing, and where it passes through the $x$-axis. Your graph does not need to be perfect, but should reflect the qualities asked above.

7-27. Consider the following graphs:




(a) Which graphs satisfy $f''(x) < 0$ when $x > 1$?

(b) Which graphs satisfy $f'(x) > 0$ when $0 < x < 1$?

(c) If $f$ is shown in Graph 3, which graph is the derivative of $f$?

7-28. Solve each of the following problems:

(a) Find the point on the line $y = 2x - 3$ closest to the origin.

(b) Find two numbers whose difference is 100 and whose product is minimal.

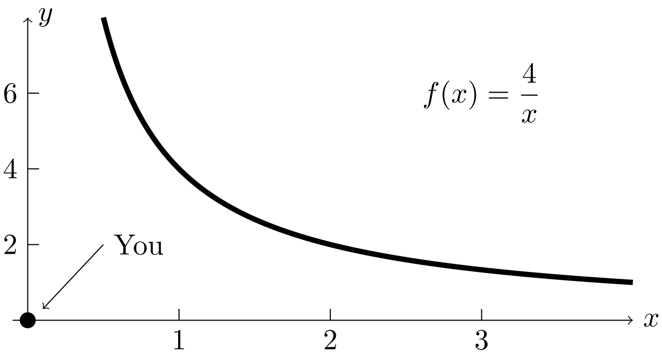
(c) A box with a square base and open top must have a volume of $32,000$ cm$^3$. Find the dimensions of the box which minimize the amount of material to be used.

(d) The top and bottom margins of a poster are each 6 cm and the side margins are 4 cm. If the area of the printed material is fixed at 384 cm$^2$, find the dimensions of the poster with the smallest area.

(e) Consider the collection of rectangles with a vertex at $(1, 0)$ and the other lying on $y = cx^2$ for some $c > 0$ and $0 \leq x \leq 1$. Find the rectangle with maximal area.

7-29. Find the point on the line $y = 3x + 5$ which is closest to the origin $(0, 0)$, and determine the corresponding distance. Be sure to show your answer is a minimum. Your solution must use calculus.

7-30. You are standing at the origin $(0, 0)$ watching a speed skating race. Your favourite competitor, Carl Friedrich Gauss, is fast approaching. From where you are standing, he takes the path $f(x) = 4/x$ for $0.5 \leq x \leq 4$. What is the shortest distance that Carl comes from you?




*Hint:* The distance between the origin $(0, 0)$ and a point $(x, y)$ is $D(x, y) = \sqrt{x^2 + y^2}$.

*Note:* The graph to the left is **not** drawn to scale.

7-31. Sketch the following functions.

(a) $f(x) = x^4 - 8x^2 + 4$

(b) $f(x) = \dfrac{1 + x^2}{1 - x^2}$

(c) $f(x) = \dfrac{1}{x^3 - x}$

(d) $f(x) = \sqrt{x} - \sqrt{x - 1}$

(e) $f(x) = e^x - x$.

7-32. For each given function $f$, find the $n$th order polynomial approximation at $x = a$.

(a) $f(x) = x^2 + 2x + 1, n = 1, a = 0$

(b) $f(x) = x^2 + 2x + 1, n = 2, a = 0$

(c) $f(x) = x^2 + 2x + 1, n = 2, a = -2$

(d) $f(x) = \ln(x^2 + 1), n = 2, a = 0$

(e) $f(x) = \sqrt{x^2 + x + 1}, n = 1, a = 1$

(f) $f(x) = e^x + e^{-x}, n = 4, a = 0$

(g) $f(x) = 1/x, n = 3, a = 1$

7-33. Use a first order approximation to determine each given value. Use a calculator to determine the error in your approximation.

(a) $\sqrt{3.9}$

(b) $\ln(1.3)$

(c) $e^{0.25}$

7-34. Repeat Exercise 7-32, but this time use a second order approximation.

7-35. Find the general nth order polynomial approximation to each function at $a = 0$.

(a) $f(x) = e^x$.

(b) $f(x) = \ln(1 + x)$

(c) $f(x) = \dfrac{1}{1 - x}$

# 8 Integration

The over-arching goal of integration is to add things together in a continuous fashion. This manifests in applications such as finding the area under a curve, or the volume of an object. In application, it's used to calculate physical quantities such as work, flux, or voltage potentials, or economic quantities such as surpluses. The fact that this is even remotely related to the process of differentiation is not at all obvious, though we will see shortly that there is in fact an intimate relationship.

## 8.1 The Definite Integral

The integral is an incredibly complex piece of mathematics, so like derivatives we will black box the majority of the process.

**Definition 8.1**

If $f$ is a function defined at every point in an interval $[a, b]$, the *integral of $f$ on* $[a, b]$ is the signed area between the $x$-axis and the graph of $f$, and is denoted

$$\int_a^b f(x)\,\mathrm{d}x.$$

By signed area, we mean that any area above the $x$-axis is given a positive weight, while area beneath the $x$-axis is assigned a negative weight. Hence it's possible for the area under a curve to cancel with itself.

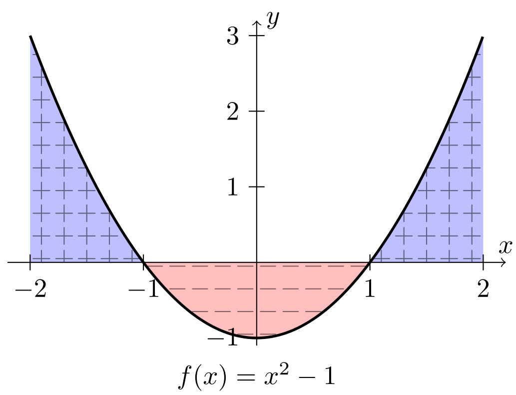



Figure 8.1: The integral computes signed area, so that area above the $x$-axis is positive, while area beneath the $x$-axis is negative.

There are a few functions whose integrals are relatively straightforward to compute.

**Example 8.2**

Let $f$ be the constant function $f(x) = c$ for some positive constant $c \in \mathbb{R}$. If $[a, b]$ is an interval, determine $\displaystyle\int_a^b f(x)\,\mathrm{d}x$.

*Solution.* The integral is the area under the curve, but $f(x) = c$ is just a straight line. The area under the curve is thus a rectangle, with height $c$ and width $b - a$. We conclude that

$$\int_a^b c\,\mathrm{d}x = c(b-a). \qquad \blacksquare$$

**Example 8.3**

Suppose that $0 < a < b$. Find the value of $\displaystyle\int_a^b x\,\mathrm{d}x$.

*Solution.* Plotting the function $f(x) = x$, the area under the graph on both $[0, a]$ and $[0, b]$ is a rectangle. The former rectangle has a height of $a$ and width of $a$, while the later has a height and width of $b$, so we can immediately conclude that

$$\int_0^a x\,\mathrm{d}x = \frac{1}{2}a^2 \quad \text{and} \quad \int_0^b x\,\mathrm{d}x = \frac{1}{2}b^2.$$

Now the question wants to know the area on the interval $[a, b]$, but this can be determined by subtracting these two rectangles, thus

$$\int_a^b x\,\mathrm{d}x = \frac{1}{2}(b^2 - a^2). \qquad \blacksquare$$

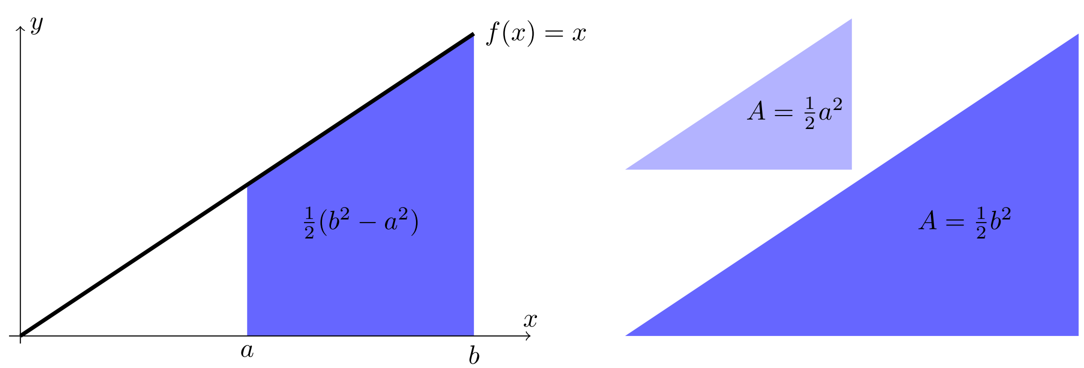



Figure 8.2: One can determine $\int_a^b x\,\mathrm{d}x$ by recognizing that the area is just a difference of triangles.

**Example 8.4**

Determine $\displaystyle\int_0^1 \sqrt{1-x^2}\,\mathrm{d}x$.

*Solution.* This one is a bit trickier. If $y = \sqrt{1-x^2}$, then we can re-arrange this to get $x^2 + y^2 = 1$; namely, that the graph of $f(x) = \sqrt{1-x^2}$ is the top half of the unit circle. Moreover, we've only been asked to determine the area on $[0,1]$, so we're looking at only the quarter circle in the first quadrant. Since a circle with radius $r=1$ has an area $\pi r^2 = \pi$, we conclude that

$$\int_0^1 \sqrt{1-x^2}\,\mathrm{d}x = \frac{\pi}{4}. \qquad \blacksquare$$

These three examples are, in some sense, the simplest. That should be intimidating, as they got difficult quickly. Indeed, much of the motivation for integral calculus comes from the fact that computing areas is *really* hard. For example, can you even guess what the area under $f(x) = x^2$ is on $[1,4]$? To develop the tools for integration, we'll have to indulge a brief intermezzo.

## 8.2 Anti-Derivatives

Before continuing, we must take a slight detour and examine the topic of anti-differentiation. Anti-differentiation is the reverse process of differentiation; that is, if I give you a function $f$ then our goal is to find a function $F$ such that $F' = f$. To this end, we have the formal definition:

**Definition 8.5**

Given a function $f$ on $[a,b]$, we say that a function $F$ is an *anti-derivative* of $f$ if $F' = f$ for all $x \in [a,b]$.

**Example 8.6**

Compute an anti-derivative of $f(x) = 5x^4$.

*Solution.* We know that polynomials differentiate to give polynomials, so let's assume that $F(x) = x^n$ for some $n$. For $F$ to be an anti-derivative of $f$ it must be that $F'(x) = nx^{n-1} = f(x) = 5x^4$. It is not too hard to see that $n = 5$ works, so that the anti-derivative of $f(x) = 5x^4$ is $F(x) = x^5$. $\blacksquare$

The previous example was exceptionally easy to solve because of the coefficient 5 in the monomial term. If that term had not been there, then we would just artificially add it. For example, the anti-derivative of $3x^4$ may be computed by realizing that

$$3x^4 = 3 \cdot \frac{5}{5}x^4 = \frac{3}{5}5x^4.$$

Since scalar multiples pass through derivatives, we hypothesize that the anti-derivative of $3x^4$ is $3x^5/5$ and a quick computation confirms this.

In fact, just using the properties of differentiation, we can immediately infer a few results about anti-derivatives. Since the derivative is linear, we have

$$\frac{\mathrm{d}}{\mathrm{d}x}cf(x) = cf'(x), \qquad \frac{\mathrm{d}}{\mathrm{d}x}\left[f(x)+g(x)\right] = f'(x)+g'(x)$$

and this tells us that anti-differentiation will also be linear. To see this, let $F$ and $G$ be anti-derivatives of $f$ and $g$, so that

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[cF(x)\right] = cF'(x) = cf(x) = c\frac{\mathrm{d}}{\mathrm{d}x}F(x)$$
$$\frac{\mathrm{d}}{\mathrm{d}x}\left[F(x)+G(x)\right] = F'(x)+G'(x) = f(x)+g(x) = \frac{\mathrm{d}}{\mathrm{d}x}F(x)+\frac{\mathrm{d}}{\mathrm{d}x}G(x).$$

Note that the anti-derivative of a function is not unique, since we may add any constant to a function to find a new anti-derivative. For example, assume that $F$ is an anti-derivative for $f$ so that $F' = f$. Define a new function $F_c(x) = F(x) + c$ for any constant $c \in \mathbb{R}$. We then have that

$$\frac{\mathrm{d}}{\mathrm{d}x}F_c(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left[F(x)+c\right] = F'(x) = f(x).$$

so that $F_c$ is also an anti-derivative. This implies that there are an entire real number's worth of functions which are the anti-derivative of a function. More concretely, Example 8.6 shows that $x^5$ is the anti-derivative of $5x^4$, but a quick computation easily shows that $x^5 + c$ also differentiates to $5x^4$ for any constant $c$.

**Corollary 8.7**

If $f$ is a function with an anti-derivative $F$, then $F$ is unique up to an additive constant; that is, if $\tilde{F}$ is any other anti-derivative of $f$, then there exists some constant $c$ such that $F(x) = \tilde{F}(x) + c$.

**Example 8.8**

Determine an anti-derivative of the function $f(x) = 3\sqrt{x}$.

*Solution.* Working backwards, we know from the Power Rule that the anti-derivative will look something like $F(x) = Ax^{3/2}$ for some unknown value of $A$. Differentiating $F$ gives

$$F'(x) = \frac{3A}{2}x^{1/2} = \frac{3A}{2}\sqrt{x}.$$

For $F$ to be an anti-derivative of $f$, we need $3A/2 = 3$ or $A = 2$, hence $F(x) = 2x^{3/2}$ is an anti-derivative. ■

**Example 8.9**

Determine an anti-derivative for $f(x) = 2^x + 12x^2$.

*Solution.* The anti-derivative of the sum is the sum of the anti-derivatives. I'll leave it as an exercise to show that an anti-derivative of $12x^2$ is $4x^3$, so we need only determine an anti-derivative for $2^x$. Let $g(x) = 2^x$, in which case an anti-derivative should look something like $G(x) = A2^x$. Differentiating $G$ gives $G'(x) = A\ln(2)2^x = g(x) = 2^x$, meaning $A\ln(2) = 1$ or $A = 1/\ln(2)$. Thus an anti-derivative for $f$ is

$$F(x) = \frac{2^x}{\ln(2)} + 4x^3.$$ ■

For reference sake, the following is a list of simple anti-derivatives where the additive constant is taken to be zero:

| Function | Anti-derivative |
|---|---|
| $x^n (n \neq -1)$ | $\dfrac{x^{n+1}}{n+1}$ |
| $\dfrac{1}{x}$ | $\ln\|x\|$ |
| $e^x$ | $e^x$ |
| $a^x$ | $\dfrac{a^x}{\ln(a)}$ |

**Example 8.10**

Compute $f$ if $f''(x) = \sqrt{x} + e^x$.

*Solution.* Notice that the second derivative is given, so we will have to compute the anti-derivative twice. Here in particular it is essential to recall that anti-derivatives are only defined up to additive constants. According to our table above, we have the following derivative - anti-derivative pairs:

$$\sqrt{x} = \frac{\mathrm{d}}{\mathrm{d}x}\frac{2}{3}x^{3/2}, \qquad e^x = \frac{\mathrm{d}}{\mathrm{d}x}e^x$$

so that the first derivative (given by the anti-derivative of $f''$) is

$$f'(x) = \frac{2}{3}x^{3/2} + e^x + c$$

for some constant $c$. It is important to include the $c$ here since when we take another anti-derivative, it will contribute to the solution. Once again, the anti-derivatives are given by

$$x^{3/2} = \frac{\mathrm{d}}{\mathrm{d}x}\frac{2}{5}x^{5/2}, \qquad e^x = \frac{\mathrm{d}}{\mathrm{d}x}e^x, \qquad c = \frac{\mathrm{d}}{\mathrm{d}x}cx,$$

so that $f$ is

$$f(x) = \frac{4}{15}x^{5/2} + e^x + cx + d$$

where $c, d$ are constants. ■

### 8.2.1 Initial Value Problems

If additional criteria are supplied, such as the value of $f$ (or its derivatives) at particular points, then a truly unique solution may be identified.

**Example 8.11**

Suppose $f'(x) = x^2 + x$ and $f(6) = 12$. Determine $f$.

*Solution.* A general anti-derivative of $f$ is of the form $F(x) = x^3/3 + x^2/2 + C$. Our goal is to determine the value of $C$ satisfying $f(6) = 12$, thus uniquely identifying $f$. Indeed,

$$f(6) = 72 + 18 + C = 80 + C = 12$$

showing that $C = 68$. Thus $f(x) = x^3/3 + x^2/2 + 68$. ∎

**Example 8.12**

Using your solution to Example 8.10, compute the unique anti-derivative which satisfies $f(0) = 10$ and $f'(0) = 0$.

*Solution.* Our above example showed that $f'(x) = 2x^{3/2}/3 + e^x + c$. By substituting $x = 0$ into this we get

$$0 = f'(0) = \frac{2}{3}0^{3/2} + e^0 + c = 1 + c$$

so that $c = -1$. Thus $f(x) = 4x^{5/2}/15 + e^x - x + d$. Substituting $x = 0$ into this gives

$$10 = f(0) = \frac{4}{15}0^{5/2} + e^0 + d = 1 + d$$

so that $d = 9$. In conclusion, the corresponding $f(x)$ is

$$f(x) = \frac{4}{15}x^{5/2} + e^x - x + 9.$$ ∎

Notice that Example 8.12 required two conditions to specify the number of constants. In general, if one is given the $n$th derivative of a function, one needs to specify $n$-conditions to uniquely determine the function.

## 8.3 The Fundamental Theorem of Calculus

In this section, we will make the connection between the theory of integration and the theory of differentiation, by means of the *Fundamental Theorem of Calculus*. Let $f$ be an integrable function on $[a, b]$, and fix the left endpoint at $a$. Now for each $x \in [a, b]$, we have an integrable function on $[a, x]$ and hence the definite integral exists and produces a number. Thus we have a function

$$F(x) = \int_a^x f(s)\,\mathrm{d}s$$

which assigns to each point $x$ the value of the definite integral on $[a, x]$. Analogous to differentiation, wherein we had a function $f$ on $[a, b]$ and created a function $f'$ on $[a, b]$ with interesting properties, we now have the function $F$ on $[a, b]$, and we are interested in its properties.

**Theorem 8.13: Fundamental Theorem of Calculus**

1. If $f$ is integrable on $[a,b]$ then $F(x) = \int_a^x f(s)\,\mathrm{d}s$ is continuous on $[a,b]$. Moreover, $F$ is differentiable at any point where $f$ is continuous, and in this case $F$ is an anti-derivative of $f$.
2. If $f$ is integrable on $[a,b]$, and $F$ be a continuous anti-derivative of $f$ which is differentiable at all but finitely many points, then
$$\int_a^b f(s)\,\mathrm{d}s = F(b) - F(a). \tag{8.1}$$

**Remark 8.14**

1. Effectively, the Fundamental Theorem of Calculus indicates that differentiation and integration are 'inverses' of one another. This is not exactly true, as Example 8.17 demonstrates.
2. The choice of anti-derivative $F$ in Theorem 8.13(2) does not matter. If $\tilde{F}$ is another anti-derivative of $f$, then by Corollary 8.7 there exists some real number $C$ such that $F(x) = \tilde{F}(x) + C$. Substituting this into (8.1) yields
$$\int_a^b f(s)\,\mathrm{d}s = F(b) - F(a) = \left[\tilde{F}(b) + C\right] - \left[\tilde{F}(a) + C\right] = \tilde{F}(b) - \tilde{F}(a).$$
3. The lower bound of integration does not matter, so long as the function stays integrable on $[a,x]$. Indeed, if $c$ is any other point such that $f$ is integrable on $[c,x]$ then
$$\int_c^x f(s)\,\mathrm{d}s = \int_a^c f(s)\,\mathrm{d}s + \int_a^x f(s)\,\mathrm{d}s = \int_a^x f(s)\,\mathrm{d}s + C$$
where $C$ is the value of the integral on $[a,c]$. Hence $\int_c^x f(s)\,\mathrm{d}s$ only differs from $F$ by an additive constant, and hence is an anti-derivative as well.

**Example 8.15**

Verify Example 8.3; that is, show that $\displaystyle\int_a^b x\,\mathrm{d}x = \frac{1}{2}(b^2 - a^2)$.

*Solution.* It suffices to find an anti-derivative of the function $f(x) = x$. The reader can quickly check that $F(x) = x^2/2$ satisfies this requirement, so by the Fundamental Theorem of Calculus:
$$\int_a^b x\,\mathrm{d}x = F(b) - F(a) = \frac{1}{2}b^2 - \frac{1}{2}a^2 = \frac{1}{2}\left(b^2 - a^2\right),$$
precisely as shown in Example 8.3. ∎

**Example 8.16**

Determine the value $\int_1^4 \frac{x+x^3}{x^4}\,\mathrm{d}x$.

*Solution.* By expanding the integrand and using linearity of the definite integral, we get

$$\int_1^4 \frac{x+x^3}{x^4}\,\mathrm{d}x = \int_1^4 \left[\frac{1}{x^3}+\frac{1}{x}\right]\mathrm{d}x = \int_1^4 \frac{1}{x^3}\,\mathrm{d}x + \int_1^4 \frac{1}{x}\,\mathrm{d}x.$$

The function $f(x) = x^{-3}$ has an anti-derivative $F(x) = -x^{-2}/2$, while $g(x) = x^{-1}$ has an anti-derivative $G(x) = \ln(x)$. By the Fundamental Theorem of Calculus, we thus have

$$\begin{aligned}\int_1^4 \frac{x+x^3}{x^4}\,\mathrm{d}x &= \int_1^4 \frac{1}{x^3}\,\mathrm{d}x + \int_1^4 \frac{1}{x}\,\mathrm{d}x = [F(4)-F(1)] + [G(4)-G(1)]\\ &= -\frac{1}{2}\left[\frac{1}{16}-1\right] + [\ln(4)-\ln(1)] = \frac{15}{32}+\ln(4).\end{aligned}$$

■

**Example 8.17**

Let $f$ be a continuous function on $\mathbb{R}$. Evaluate

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_0^x f(t)\,\mathrm{d}t - \int_0^x \frac{\mathrm{d}}{\mathrm{d}t}f(t)\,\mathrm{d}t.$$

*Solution.* If integration and differentiation were truly inverses, then this would simply evaluate to zero. However, let us be a bit more prudent in our evaluation. By the Fundamental Theorem of Calculus, $F(x) = \int_0^x f(t)\,\mathrm{d}t$ is an anti-derivative of $f(x)$, and hence

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_0^x f(t)\,\mathrm{d}t = f(x).$$

On the other hand, $f$ is clearly an anti-derivative of $f'$, and so

$$\int_0^x \frac{\mathrm{d}}{\mathrm{d}t}f(t)\,\mathrm{d}t = \int_0^x f'(t)\,\mathrm{d}t = f(x) - f(0).$$

Hence the difference between these two terms comes out to $f(0)$; that is, they differ up to a constant. ■

**Example 8.18**

Determine $G'(x)$ if $G(x) = \int_{-1}^{x^2} te^t\,\mathrm{d}t$.

*Solution.* You should stare at this equation until you are convinced that this is a function of $x$. To proceed, define a function

$$F(x) = \int_{-1}^{x} xe^x \, \mathrm{d}t$$

which, according to the Fundamental Theorem of Calculus, is an anti-derivative of the function $f(x) = xe^x$. We can write $G$ in terms of $F$, since

$$G(x) = \int_{-1}^{x^2} xe^x \, \mathrm{d}t = F(x^2).$$

We can thus differentiate $G$ using the Chain Rule,

$$G'(x) = \frac{\mathrm{d}}{\mathrm{d}x} F(x^2) = 2xF'(x^2) = 2xf(x^2) = 2x(x^2 e^{x^2}) = 2x^3 e^{x^2}. \qquad \blacksquare$$

### 8.3.1 Properties of the Definite Integral

The definite integral satisfies the following properties:

1. **Additivity of Domain:** If $f$ is integrable on $[a, b]$ and $[b, c]$ then $f$ is integrable on $[a, c]$ and
$$\int_a^c f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_b^c f(x)\,\mathrm{d}x.$$
2. **Additivity of Integral:** If $f, g$ are integrable on $[a, b]$ then $f + g$ is integrable on $[a, b]$ and
$$\int_a^b [f(x) + g(x)]\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x.$$
3. **Scalar Multiplication:** If $f$ is integrable on $[a, b]$ and $c \in \mathbb{R}$, then $cf$ is integrable on $[a, b]$ and
$$\int_a^b cf(x)\,\mathrm{d}x = c\int_a^b f(x)\,\mathrm{d}x.$$
4. **Inherited Integrability:** If $f$ is integrable on $[a, b]$ then $f$ is integrable on any subinterval $[c, d] \subseteq [a, b]$.
5. **Monotonicity of Integral:** If $f, g$ are integrable on $[a, b]$ and $f(x) \leq g(x)$ for all $x \in [a, b]$ then
$$\int_a^b f(x)\,\mathrm{d}x \leq \int_a^b g(x)\,\mathrm{d}x.$$
6. **Subnormality:** If $f$ is integrable on $[a, b]$ then $|f|$ is integrable on $[a, b]$ and satisfies
$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \leq \int_a^b |f(x)|\,\mathrm{d}x.$$

**Example 8.19**

Determine the value of $\displaystyle\int_0^3 \left[x+\sqrt{x}\right]\,\mathrm{d}x$.

*Solution.* Since the integral is additive, we know that

$$\int_0^4 \left[x+\sqrt{x}\right]\,\mathrm{d}x = \int_0^4 x\,\mathrm{d}x + \int_0^4 \sqrt{x}\,\mathrm{d}x.$$

Computing each of these separately we get

$$\int_0^4 x\,\mathrm{d}x = \frac{1}{2}x^2\Big|_0^4 = 8 \quad \text{and} \quad \int_0^4 \sqrt{x}\,\mathrm{d}x = \frac{2}{3}x^{3/2}\Big|_0^4 = \frac{16}{3}.$$

so that $\int_0^4 \left[x+\sqrt{x}\right]\,\mathrm{d}x = 8 + 16/3 = 40/3$. ∎

**Example 8.20**

Determine $\displaystyle\int_{-1}^1 f(x)\,\mathrm{d}x$ where $f(x) = \begin{cases} x^2 & -1 < x \leq 0 \\ x & 0 < x < 1 \end{cases}$.

*Solution.* By Additivity of Domain we can split this into the integral on $[-1,0]$ and $[0,1]$. Indeed,

$$\begin{aligned}\int_{-1}^1 f(x)\,\mathrm{d}x &= \int_{-1}^0 f(x)\,\mathrm{d}x + \int_0^1 f(x)\,\mathrm{d}x = \int_{-1}^0 x^2\,\mathrm{d}x + \int_0^1 x\,\mathrm{d}x \\ &= \left[\frac{1}{3}x^3\right]_0^1 + \left[\frac{1}{2}x^2\right]_0^1 = \frac{1}{3} + \frac{1}{2} = \frac{2}{3}.\end{aligned}$$ ∎

If $a < b$, we don't have a meaningful way of interpreting the integral $\int_b^a f(x)\,\mathrm{d}x$. To assign a meaning to this integral, we wish to preserve Additivity of Domain. First off, we demand that for any integrable function $f$,

$$\int_a^a f(x)\,\mathrm{d}x = 0,$$

as there will be no area under the graph of $f$. If Additivity of Domain holds, we should thus have

$$\int_a^b f(x)\,\mathrm{d}x + \int_b^a f(x)\,\mathrm{d}x = \int_a^a f(x)\,\mathrm{d}x = 0 \quad \text{from which} \quad \int_b^a f(x)\,\mathrm{d}x = -\int_a^b f(x)\,\mathrm{d}x.$$

When the bounds of integration are reversed, we can thus evaluate the integral by imposing a negative sign on the correctly oriented bounds.

### 8.3.2 Indefinite Integrals

We have seen that integration and differentiation are spiritual inverses of one another, up to an additive constant. In particular, for a function $f$ on $[a,b]$ we saw that $F(x) = \int_a^x f(s)\,\mathrm{d}s$ is an anti-derivative of $f$, and for any anti-derivative $G$ of $f$ we have

$$G(b) - G(a) = \int_a^b f(s)\,\mathrm{d}s.$$

Anti-derivatives are unique up to constants, so there exists some constant $C$ such that $F = G + C$, with $F$ playing a particularly nice representative. However, the constant seems rather artificial: we know that the anti-derivative of $x^3$ is $x^4/4 + C$, but the meat-and-bones lies with the $x^4/4$ term, not the constant. Hence our goal for this section is to represent the entire class of anti-derivatives, something called the *indefinite integral*.

The indefinite integral does not concern itself with upper and lower bounds of integration – our goal is to represent an entire class of functions, and imposing bounds forces us to look at particular representatives. Consequently, we denote the indefinite integral with the usual integral sign, albeit with the bounds omitted:

$$\int f(x)\,\mathrm{d}x.$$

Remember, this notation means *the entire set of anti-derivatives.*

**Example 8.21**

Determine the following indefinite integrals:

1. $\displaystyle\int \left(\frac{x^4 + 2x^2 + 1}{x^3}\right) \mathrm{d}x$,
2. $\displaystyle\int f(x)f'(x)\,\mathrm{d}x$, where $f$ is differentiable.

*Solution.* In time, we will learn more systematic ways of determining these integrals, but for now we will need to use the clever part of our brains to find appropriate classes of anti-derivatives.

1. Notice that we can re-write the integrand as
$$\frac{x^4 + 2x^2 + 1}{x^3} = x + \frac{2}{x} + \frac{1}{x^3}.$$
We are well acquainted with the functions which yield these as derivatives, and we get
$$\int \frac{x^4 + 2x^2 + 1}{x^3} = \int \left(x + \frac{2}{x} + \frac{1}{x^3}\right) \mathrm{d}x = \frac{1}{2}x^2 + 2\ln(x) - \frac{1}{2x^2} + C.$$

2. This problem is a little more abstract: We need to find a function which differentiates to $f(x)f'(x)$. If we think hard, we see that $\frac{\mathrm{d}}{\mathrm{d}x}[f(x)]^2 = 2f(x)f'(x)$, so by dividing by 2 we will get the desired integrand. Applying the Fundamental Theorem of Calculus, we thus get
$$\int f(x)f'(x)\,\mathrm{d}x = \int \frac{\mathrm{d}}{\mathrm{d}x} f(x)^2\,\mathrm{d}x = f(x)^2 + C. \tag{8.2}$$
We will see more on how to solve integral like this in Section 8.4.1. ■

### 8.3.3 Integral Notation

We have been dramatically overloading our use of the integral sign. If $f$ is integrable on $[a,b]$, we have seen three different objects

$$\int_a^b f(s)\,\mathrm{d}s, \quad F(x) = \int_a^x f(s)\,\mathrm{d}s, \quad \text{and} \quad \int f(s)\,\mathrm{d}s.$$

The first is simply a *number* which represents the signed area under the function $f$, the second is a *function* which assigns to each $x$ the area under the function from $a$ to $x$, the third is an infinite *family of functions* representing all anti-derivatives of $f$. They are intimately related to be certain, but each has a very different lifestyle. One must be careful not to confuse the relationships.

Additionally, some authors prefer to the use the notation

$$\int \mathrm{d}s f(s) \quad \text{instead of} \quad \int f(s)\,\mathrm{d}s.$$

There are occasions when this is useful, but I will never use this notation in these notes.

## 8.4 Integration Techniques

In the following sections, we will develop a plethora of tools to help us compute integrals.

### 8.4.1 Integration by Substitution

Having seen that integration and differentiation are essentially inverses, we would like to develop some techniques and rules form computing integrals. It should be unsurprising that those rules will arise as the "inverse" operations of the rules obtained from differential calculus. We recall the chain rule of differential calculus tells us that

$$\frac{\mathrm{d}}{\mathrm{d}x} f(g(x)) = f'(g(x))g'(x)$$

hence by applying the Fundamental Theorem of Calculus, we see that

$$\int f'(g(x))g'(x)\,\mathrm{d}x = f(g(x)) + C. \tag{8.3}$$

Unfortunately, the majority of times nature will conspire against us and not write our integrand so plainly as $f'(g(x))g'(x)$. Hence we develop some techniques to make life simpler. Our strategy is as follows:

1. Look to see if we can find the occurrence of a function and its derivative. In (8.3) above, we are looking for the function $g$, since it occurs in the argument of $f$ and its derivative appears as $g'$.
2. Define a new variable, $u = g(x)$ so that $\frac{\mathrm{d}u}{\mathrm{d}x} = g'(x)$. We often write this second equation as $\mathrm{d}u = g'(x)\,\mathrm{d}x$.

3. Replace all $x$ dependencies with $u$ dependencies. Namely, recognize that

$$\underbrace{f'(g(x))}_{f(u)}\underbrace{g'(x)\,\mathrm{d}x}_{\mathrm{d}u} = f'(u)\,\mathrm{d}u$$

4. Using the Fundamental Theorem of Calculus, evaluate our new integral

$$\int f'(u)\,\mathrm{d}u = f(u) + C.$$

5. We now have our solution, but it is in terms of the variable $u$. This is not a problem since we know that $u = g(x)$, so we just make this substitution to get our final solution

$$\int f'(g(x))g'(x)\,\mathrm{d}x = f(g(x)) + C.$$

Let's try a simple example:

**Example 8.22**

Determine $\int 2xe^{x^2}\,\mathrm{d}x$.

*Solution.* Here we see a function $g(x) = x^2$ and its derivative $g'(x) = 2x$, so we make the substitution $u = x^2$, $\mathrm{d}u = 2x\,\mathrm{d}x$. Substituting everything back in, we get

$$\int e^{x^2}\underbrace{2x\,\mathrm{d}x}_{\mathrm{d}u} = \int e^u\,\mathrm{d}u = e^u + C = e^{x^2} + C.$$

We can check that our answer is correct by differentiating. ■

It was convenient that the integrand included the 2 necessary to make this substitution work. On the other hand, since anti-derivatives don't care about scalar multiples, we should still be able to handle the above integral if the 2 were replaced with something different. There are two paradigms for how to attack this. Say for example we wanted to integrate

$$\int 10xe^{x^2}\,\mathrm{d}x.$$

Recognizing that $x^2$ and $10x$ are very close matches, we could artificially introduce the $2x$ necessary to make the substitution work. For example, by writing

$$\int 10xe^{x^2}\,\mathrm{d}x = 5\int 2xe^{x^2}\,\mathrm{d}x = 5e^{x^2} + C.$$

Or, we can make the manipulation in the differentials. For example, we would still set $u = x^2$ so that $\mathrm{d}u = 2x\,\mathrm{d}x$. Our integral still has an $x\,\mathrm{d}x$ term, so we solve for it to get $x\,\mathrm{d}x = \mathrm{d}u/2$. Substituting this into the integral,

$$\int 10e^{x^2}\underbrace{x\,\mathrm{d}x}_{\mathrm{d}u/2} = \int \frac{10}{2}e^u\,\mathrm{d}u = 5\int e^u\,\mathrm{d}u = 5e^{x^2} + C.$$

Hence our goal, when performing a substitution, is to find the correct function/derivative pair only up to a scalar multiple.

**Example 8.23**

Determine the integral $\int x^3\sqrt{x^4+10}\,\mathrm{d}x$.

*Solution.* Here we see a function derivative pair $f(x) = x^4$ and $f'(x) = x^3$, but note there is no harm in making $f(x) = x^4 + 10$ since this doesn't affect the derivative. Setting $u = x^4 + 10$ gives $\mathrm{d}u = 4x^3\,\mathrm{d}x$, and

$$\int x^3\sqrt{x^4+10}\,\mathrm{d}x = \frac{1}{4}\int \sqrt{u}\,\mathrm{d}u = \frac{1}{6}u^{3/2} + C = \frac{1}{6}(x^4+10)^{3/2} + C. \qquad \blacksquare$$

**Example 8.24**

For $a, b \neq 0$, compute $\int \frac{x^{n-1}}{\sqrt{a+bx^n}}\,\mathrm{d}x$.

*Solution.* Following the above program, our first step should be to identify a function and its derivative. The fact that there is an $x^n$ and an $x^{n-1}$ is a pretty good sign. Since constants do not affect the integration, we can make our lives *even easier* if we define $u = a + bx^n$ so that $\mathrm{d}u = bnx^{n-1}\,\mathrm{d}x$. Unfortunately, there is no $bnx^{n-1}\,\mathrm{d}x$ in the integrand, but there is an $x^{n-1}\,\mathrm{d}x$. Since these are related only up to a constant, we can divide both sides to find that $x^{n-1}\,\mathrm{d}x = \mathrm{d}u/bn$. Adding our substitutions we then get

$$\int \frac{x^{n-1}}{\sqrt{a+bx^n}}\,\mathrm{d}x = \int \frac{\frac{1}{bn}\,\mathrm{d}u}{\sqrt{u}} = \frac{1}{bn}\int \frac{1}{\sqrt{u}}\,\mathrm{d}u.$$

This is now a very simple integral to calculate, and indeed we find that

$$\frac{1}{bn}\int \frac{1}{\sqrt{u}}\,\mathrm{d}u = \frac{2}{bn}\sqrt{u} + C.$$

We need this to be in terms of $x$ rather than $u$, so we recall that $u = a + bx^n$ to finally find that

$$\int \frac{x^{n-1}}{\sqrt{a+bx^n}}\,\mathrm{d}x = \frac{2}{bn}\sqrt{a+bx^n} + C. \qquad \blacksquare$$

Substitution is not just handy for applying the chain rule. It also allows us to "change variables."

**Example 8.25**

Compute $\int x\sqrt{x+1}\,\mathrm{d}x$.

*Solution.* Notice that if we could somehow switch the $x$ and the $x+1$, this integral would be much simpler, since then $(x+1)\sqrt{x} = x^{3/2} + x$. Normally in mathematics, if we want to do such a thing, we just define a new variable $u = x+1$ so that $x = u-1$ and then a similar trick to the one above will work.

Since we are working with an integral though, we must be a bit more careful. We shall still define $u = x+1$ with $x = u-1$, but we must also track the differentials. Luckily, in this case $\mathrm{d}u = \mathrm{d}x$ and there is nothing to do. We thus get

$$\int x\sqrt{x+1}\,\mathrm{d}x = \int (u-1)\sqrt{u}\,\mathrm{d}u = \int (u^{3/2} - \sqrt{u})\,\mathrm{d}u = \frac{2}{5}u^{5/2} - \frac{2}{3}u^{3/2} + C.$$

Converting back to a function of $x$ yields

$$\int x\sqrt{x+1}\,\mathrm{d}x = \frac{2}{5}(x+1)^{5/2} - \frac{1}{2}(x+1)^2 + C. \qquad \blacksquare$$

**Definite Integrals** When dealing with definite integrals, we adhere to the same process as indefinite integrals, but we must also accommodate the lower and upper bounds of integration. Let $f$ be a continuous function with anti-derivative $F$, while $g$ is continuously differentiable. Once again consider the case when we are integrating the function

$$\int_a^b f(g(x))g'(x)\,\mathrm{d}x.$$

We know that $F(g(x))$ is an integrable anti-derivative, so the Fundamental Theorem of Calculus implies that

$$\int_a^b f(g(x))g'(x)\,\mathrm{d}x = [F(g(x))]_a^b = F(g(b)) - F(g(a)).$$

This implies that the correct lower and upper bounds of integration are $g(a)$ and $g(b)$ respectively, since

$$\int_a^b f(g(x))g'(x)\,\mathrm{d}x = \int_{g(a)}^{g(b)} f(u)\,\mathrm{d}u = [F(u)]_{g(a)}^{g(b)} = F(g(b)) - F(g(a))$$

gives the correct solution.

For a different perspective, the lower and upper bounds of integration say that our variable $x$ is moving between the values $x = a$ and $x = b$. If we make the substitution $u = g(x)$, then our new variable is $u$. As $x$ goes from $a$ to $b$, then $u$ goes between $g(a)$ and $g(b)$, and so our integral becomes

$$\int_a^b f'(g(x))g'(x)\,\mathrm{d}x = \int_{g(a)}^{g(b)} f'(u)\,\mathrm{d}u.$$

As an example, consider what happens if we **fail to change the bounds of integration.** Consider the integral

$$\int_1^3 [x-2]\,\mathrm{d}x = \left[\frac{1}{2}x^2 - 2x\right]_1^3 = \left(\frac{9}{2} - 6\right) - \left(\frac{1}{2} - 2\right) = -\frac{3}{2} - (-\frac{3}{2}) = 0.$$

On the other hand, if we make the substitution $u = x - 2$ with $\mathrm{d}u = \mathrm{d}x$, then

$$\int_1^3 u \, \mathrm{d}u = \frac{1}{2}u^2 \Big|_1^3 = \frac{9}{2} - \frac{1}{2} = 4.$$

This is a different answer! Indeed, think about the plots of $y = x$ and $y = x - 2$. For the integrals to be the same, we need to change the corresponding domain.

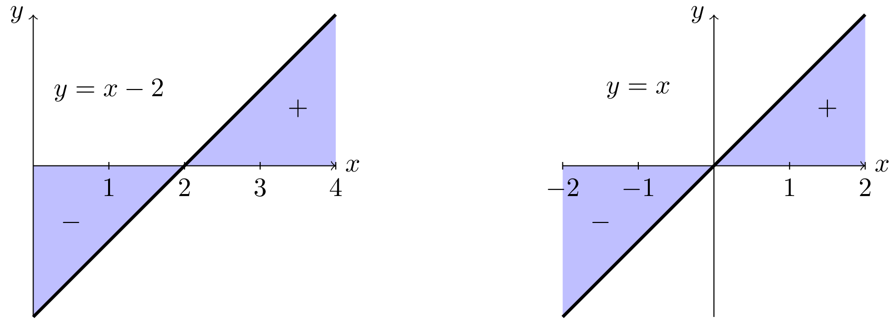



Figure 8.3: When we make a substitution, we must change the domain to accommodate the fact that we've changed the integrand.

**Example 8.26**

Determine the value of $\displaystyle\int_0^1 x\sqrt{2+x^2}\, \mathrm{d}x$.

*Solution.* We will proceed using the substitution $u = 2 + x^2$ so that $\mathrm{d}u = 2x\, \mathrm{d}x$. When $x = 0$ we have $u = 2$, while when $x = 1$ we get $u = 3$, so

$$\int_0^1 x\sqrt{2+x^2}\, \mathrm{d}x = \frac{1}{2}\int_2^3 \sqrt{u}\, \mathrm{d}u = \frac{1}{3}u^{3/2}\Big|_{u=2}^{u=3} = \frac{3\sqrt{3} - 2\sqrt{2}}{3}. \qquad \blacksquare$$

Note that when performing the definite integral, we do not need to convert back to the $x$-representation, since our upper and lower bounds have already accommodated for that change.

**Example 8.27**

Determine the integral $\displaystyle\int_2^4 \frac{1}{x\ln(2x)}\, \mathrm{d}x$.

*Solution.* We will proceed using the substitution $u = \ln(2x)$ so that $\mathrm{d}u = x^{-1}\, \mathrm{d}x$. When $x = 2$ we have $u = \ln(4)$, while when $x = 4$ we have $u = \ln(8)$, so that

$$\int_2^4 \frac{1}{x\ln(2x)}\, \mathrm{d}x = \int_{\ln(4)}^{\ln(8)} \frac{1}{u}\, \mathrm{d}u = \ln(u)\Big|_{\ln(4)}^{\ln(8)} = \ln(\ln(8)) - \ln(\ln(4)). \qquad \blacksquare$$

### 8.4.2 Integration by Parts

Just as Integration by Substitution was the integral version of the chain rule, Integration by Parts is the analog of the product rule. Namely, we know that if $u$ and $v$ are functions, then

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[u(x)\cdot v(x)\right] = u'(x)v(x) + u(x)v'(x).$$

Integrating and applying the Fundamental Theorem of Calculus, we then find that

$$\begin{aligned}\int \frac{\mathrm{d}}{\mathrm{d}x}(u\cdot v)\,\mathrm{d}x = u(x)\cdot v(x)&\\ &= \int v\frac{\mathrm{d}u}{\mathrm{d}x}\,\mathrm{d}x + \int u\frac{\mathrm{d}v}{\mathrm{d}x}\,\mathrm{d}x\\ &= \int v\,\mathrm{d}u + \int u\,\mathrm{d}v\end{aligned}$$

which we may re-arrange to find

$$\int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u. \tag{8.4}$$

In the event of the definite integral, the bounds of integration can be carried throughout; that is,

$$\int_a^b u\,\mathrm{d}v = uv\Big|_a^b - \int_a^b v\,\mathrm{d}u.$$

Our strategy should be as follows: Assume we are integrating a product $\int f(x)g(x)\,\mathrm{d}x$. We want to choose a candidate for $\mathrm{d}v$ and for $u$. Since we will need to integrate $\mathrm{d}v$, it is often best to choose a function which is easy to integrate:

1. First look at the integrand and see if we can apply substitution. If so, do not worry about integration by parts.
2. Choose $\mathrm{d}v$ and $u$ (I often choose $\mathrm{d}v$ to be whichever function is easiest to integrate),
3. Compute $v$ by integrating $\int \mathrm{d}v = \int f(x)\,\mathrm{d}x$. Compute $\mathrm{d}u$ by differentiating $u$.
4. Substitute all appropriate variables into (8.4).

This is just a general idea of how you should proceed. To give some insight as to what is happening, consider Equation (8.4) by omitting the $uv$-term:

$$\int u(x)v'(x)\,\mathrm{d}x = -\int u'(x)v(x)\,\mathrm{d}x. \tag{8.5}$$

This is the power of integration by parts: It effectively allows us to transfer the derivative from one function to another!

**Example 8.28**

Evaluate the integral $\displaystyle\int x\ln(x)\,\mathrm{d}x$.

*Solution.* Set $u = \ln(x)$ and $\mathrm{d}v = x\,\mathrm{d}x$. Computing $\mathrm{d}u$ and $v$ we find that

$$\begin{array}{ll} u = \ln(x) & \mathrm{d}v = x\,\mathrm{d}x \\ \mathrm{d}u = \mathrm{d}x/x & v = x^2/2 \end{array}.$$

Plugging these into (8.4)

$$\begin{aligned} \int x\ln(x)\,\mathrm{d}x &= \frac{x^2\ln(x)}{2} - \int \frac{x^2}{2}\frac{1}{x}\,\mathrm{d}x = \frac{x^2\ln(x)}{2} - \frac{1}{2}\int x\,\mathrm{d}x \\ &= \frac{x^2\ln(x)}{2} - \frac{1}{4}x^2 + C. \end{aligned}$$

The best thing about integration is that you can always check your answers by differentiating. Try it! ■

**Example 8.29**

Determine $\displaystyle\int_0^3 xe^x\,\mathrm{d}x$.

*Solution.* We will take as our integration by parts

$$\begin{array}{ll} u = x & \mathrm{d}v = e^x\,\mathrm{d}x \\ \mathrm{d}u = \mathrm{d}x & v = e^x \end{array},$$

which gives us the integral

$$\int_0^3 xe^x\,\mathrm{d}x = xe^x\Big|_0^3 - \int_0^3 e^x\,\mathrm{d}x = 3e^3 - \Big[e^x\Big]_0^3 = 2e^3 + 1.$$ ■

Alternatively, there are time when the integrand does not look like a product, but we may still apply integration by parts.

**Example 8.30**

Compute $\displaystyle\int \ln(x)\,\mathrm{d}x$.

*Solution.* Looking at the integrand, there do not immediately appear to be two functions, so how can we apply integration by parts? The solution is to realize that $\ln(x) = \ln(x)\cdot 1$, so that the constant function $1(x) = 1$ is actually our second function. I think that 1 is really easy to integrate, so let us set $\mathrm{d}v = 1\cdot\mathrm{d}x$ and $u = \ln(x)$ to find that

$$\begin{array}{ll} u = \ln(x) & \mathrm{d}v = 1\cdot\mathrm{d}x \\ \mathrm{d}u = \mathrm{d}x/x & v = x \end{array}.$$

Substituting these values into (8.4) we find

$$\begin{aligned}\int \ln(x)\, \mathrm{d}x &= x\ln(x) - \int \frac{x}{x}\, \mathrm{d}x \\ &= x\ln(x) - x + C.\end{aligned}$$

Again, try differentiating this to ensure that it works! ■

**Remark 8.31** Why can we use Integration by Parts on $\int \ln(x)\, \mathrm{d}x$, and to what other functions does this same trick apply? As an exercise, you can use Integration by Parts to show that if $f$ is an invertible, integrable function with anti-derivative $F$ and inverse $f^{-1}$, then

$$\int f^{-1}(x)\, \mathrm{d}x = xf^{-1}(x) - F(f^{-1}(x)) + C.$$

Thus invertible, integrable functions can be integrated using integration by parts.

**Example 8.32**

Determine $\displaystyle\int x^3 e^{x^2}\, \mathrm{d}x$.

*Solution.* It might be tempting to start with substitution, but setting $u = x^3$ with $\mathrm{d}u = 3x^2\, \mathrm{d}x$ would result in a $\mathrm{d}u$ in the exponent, and that doesn't make any sense. Instead, let's use integration by parts. Were you to try $u = e^{x^2}$ and $\mathrm{d}v = x^3\, \mathrm{d}x$, the integral would not simplify. Instead, let's integrate by parts, with

$$\begin{array}{ll} u = x^2 & \mathrm{d}v = xe^{x^2}\, \mathrm{d}x \\ \mathrm{d}u = 2x\, \mathrm{d}x & v = e^{x^2}/2 \end{array}.$$

Substituting this into our formula gives

$$\int x^3 e^{x^2}\, \mathrm{d}x = \frac{x^2 e^{x^2}}{2} - \int xe^{x^2}\, \mathrm{d}x.$$

The term still under an integral can be solved by means of a simple substitution, and we find that

$$\int x^3 e^{x^2}\, \mathrm{d}x = \frac{x^2 e^{x^2} - e^{x^2}}{2} = \frac{(x^2-1)e^{x^2}}{2}.$$ ■

## 8.5 Exercises

8-1. Find a general form anti-derivative for the following functions:

(a) $f(x) = 7x^{2/5} + 8x^{-4/5}$

(b) $g(t) = \dfrac{1+t+t^2}{\sqrt{t}}$

(c) $h(x) = 3 \cdot 2^x + 2 \cdot 3^x$

(d) $r(z) = \frac{1}{z+1}$

8-2. Repeat Exercise 8-1, but now find the unique anti-derivative satisfying the initial condition below:

(a) $F(1) = 15$
(b) $G(0) = -5$
(c) $H(1) = 0$
(d) $R(0) = 4$

8-3. Find all such $f(t)$ such that $f'''(t) = e^t + t^{-4}$

8-4. (a) Find an anti-derivative of $f(x) = e^{2x}$.
(b) Find an anti-derivative of $g(x) = 2^{-x}$.
(c) Generalize the above results as follows. Suppose that $F$ is an anti-derivative of $f$. Find an anti-derivative of the new function $g(x) = f(ax)$.

8-5. Evaluate the integrals:

(a) $\displaystyle\int_1^2 \left(1 + x + x^2\right)\ \mathrm{d}x$
(b) $\displaystyle\int_0^1 (x^e + e^x)\ \mathrm{d}x$
(c) $\displaystyle\int_0^3 (2x - e^x)\ \mathrm{d}x$
(d) $\displaystyle\int_1^2 \frac{4 + u^2}{u^3}\ \mathrm{d}u$
(e) $\displaystyle\int_1^5 \frac{\sqrt{z} + z^2}{z}\ \mathrm{d}z$

8-6. Find the derivative of the functions:

(a) $h(x) = \displaystyle\int_1^{e^x} \ln(t)\ \mathrm{d}t$
(b) $f(x) = \displaystyle\int_x^{\pi} \sqrt{1 + t^2}\ \mathrm{d}t$
(c) $g(x) = \displaystyle\int_1^{\sqrt{x}} \frac{z^2}{z^4 + 1}\ \mathrm{d}z$
(d) $r(x) = \displaystyle\int_{2x}^{3x} \frac{u^2 - 1}{u^2 + 1}\ \mathrm{d}u$

8-7. If $f$ is continuous and $g$ and $h$ are differentiable functions, find a formula for $\displaystyle\frac{\mathrm{d}}{\mathrm{d}x}\int_{g(x)}^{h(x)} f(t)\ \mathrm{d}t$

8-8. Consider the function $f(x) = \dfrac{x^2 - 1}{x}$, and define the function

$$F(x) = \int_1^x f(t)\,\mathrm{d}t$$

with domain $(0, \infty)$.

(a) What are the critical points of $F$?
(b) Find the intervals of increasing and decreasing for $F$.
(c) Find the global maximum and minimum of $F$ on the interval $[1/2, 3]$.

8-9. Suppose that $f$ is a differentiable function, and takes on the following values

| $x$ | $-2$ | $-1$ | $0$ | $1$ | $2$ |
|---|---|---|---|---|---|
| $f(x)$ | $4$ | $0$ | $-2$ | $10$ | $7$ |
| $f'(x)$ | $-1$ | $2$ | $4$ | $-3$ | $1$ |

Define the function $G(x) = \int_x^{x^2} f'(t)\,\mathrm{d}x$.

(a) Determine the values of $G(-1)$ and $G(1)$.
(b) Determine the value of $G'(-1)$

8-10. Suppose $f(x) = \int_0^x \frac{t}{t^2+1}\,\mathrm{d}t$. If $g(x) = \int_1^{f(x)} te^{-t}\,\mathrm{d}t$, what is $g'(1)$?

8-11. Suppose that $f$ is a function such that $f(2) = 2$ and $f'(2) = 3$. Define $g(x) = \int_0^{f(x)} f(f(t))\,\mathrm{d}t$. Find $g'(2)$.

8-12. Evaluate the integrals:

(a) $\int (2x^4 - 3x)^5 (8x^3 - 3)\,\mathrm{d}x$
(b) $\int 3x(x^2-2)^{3/2}\,\mathrm{d}x$
(c) $\int x^2 e^{x^3}\,\mathrm{d}x$
(d) $\int \frac{(\ln x)^2}{x}\,\mathrm{d}x$

8-13. Evaluate:

(a) $\int_{-1}^1 \frac{1}{(5-z)^2}\,\mathrm{d}z$
(b) $\int_0^1 \frac{e^z+1}{e^z+z}\,\mathrm{d}z$
(c) $\int_0^1 t^3 e^{-t^4}\,\mathrm{d}t$
(d) $\int_0^a (x\sqrt{a^2+x^2})\,\mathrm{d}x$
(e) $\int_1^2 \frac{2x^2}{\sqrt{4x^3-2}}\,\mathrm{d}x$

8-14. Evaluate the following integrals:

(a) $\int 3xe^{4x}\,\mathrm{d}x$
(b) $\int x\ln(x)\,\mathrm{d}x$
(c) $\int x^3 e^{-x}\,\mathrm{d}x$
(d) $\int s2^s\,\mathrm{d}s$
(e) $\int \ln(\sqrt{z})\,\mathrm{d}z$
(f) $\int x(x^2+1)\ln(x^2+1)\,\mathrm{d}x$.

8-15. Evaluate the following integrals:

(a) $\int_1^e t^2\ln(t)\,\mathrm{d}t$
(b) $\int_{-1}^1 (3-x)e^{-x}\,\mathrm{d}x$
(c) $\int_1^4 \sqrt{t}\ln(t)\,\mathrm{d}t$
(d) $\int_1^9 e^{\sqrt{s}}\,\mathrm{d}s$
(e) $\int_0^1 x^5 e^{x^2}\,\mathrm{d}x$

8-16. Suppose you are given the definite integral $\int_a^b f(x)g(x)\,\mathrm{d}x$.

(a) Suppose that $G$ is an anti-derivative of $g$. Perform the integration by parts if $u = f(x)$ and $\mathrm{d}v = g(x)\,\mathrm{d}x$. Express your answer in terms of $f, g$, and $G$.

(b) Suppose $g(x) = f'(x)$. Perform the substitution $u = f(x)$. Express your answer entirely in terms of $f$.

8-17. (a) You are told that $f$ is a differentiable function such that $f'(x) = \dfrac{1}{x^3+1}$ with $f(0) = 0$ and $f(1) = \pi/4$. Determine the definite integral $\displaystyle\int_0^1 x f(x)\,\mathrm{d}x$.

(b) Suppose you are told that $f$ is a differentiable such that $f(0) = 2$ and $f(1) = 4$. Determine the definite integral $\displaystyle\int_0^1 \frac{f'(x)}{\sqrt{f(x)}}\,\mathrm{d}x$.

8-18. Suppose that $f$ is twice differentiable has the following table of values:

| $x$ | $-2$ | $-1$ | $0$ | $1$ | $2$ |
|---|---|---|---|---|---|
| $f(x)$ | $4$ | $0$ | $-2$ | $10$ | $7$ |
| $f'(x)$ | $-1$ | $2$ | $4$ | $-3$ | $1$ |

Determine the following integrals:

(a) $\displaystyle\int_{-1}^1 f'(x)f(x)\,\mathrm{d}x$.

(b) $\displaystyle\int_0^2 x f''(x)\,\mathrm{d}x$.

8-19. Evaluate each of the following integrals:

(a) $\displaystyle\int x^3 e^{-4x}\,\mathrm{d}x$

(b) $\displaystyle\int_1^2 x^n \ln(x)\,\mathrm{d}x$, where $n \in \mathbb{N}$.

(c) $\displaystyle\int_0^1 x^3\sqrt{1-x^2}\,\mathrm{d}x$

(d) $\displaystyle\int \frac{x^3}{1+x^2}\,\mathrm{d}x$

8-20. Recall from Exercise 6-19 the following two functions:

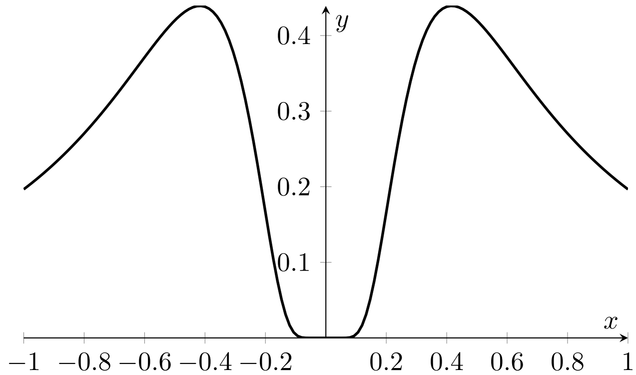
$$f(x) = \frac{e^x + e^{-x}}{2} \quad \text{and} \quad g(x) = \frac{e^x - e^{-x}}{2}.$$

You showed that $f'(x) = g(x)$, $g'(x) = f(x)$, and $f(x)^2 - g(x)^2 = 1$. Use this to determine the integral $\displaystyle\int f(x)^4 g(x)^5\,\mathrm{d}x$.

8-21. Consider the function

$$f(x) = \begin{cases} \dfrac{e^{1/x}}{x^2(e^{1/x}+1)^2} & x \neq 0 \\ 0 & x = 0 \end{cases}.$$

It is possible to show that $\lim_{x\to 0} f(x) = 0$, so that $f$ is everywhere continuous.



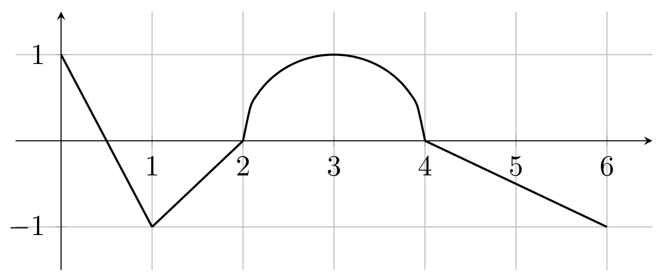

(a) Determine $\int f(x)\,\mathrm{d}x$.

(b) Using your graphing software of choice, plot your solution from part (a). Quickly sketch the plot you get. Explain why this does not contradict the Fundamental Theorem of Calculus.

8-22. Suppose $f$ is integrable and $\int_0^4 f(x)\,\mathrm{d}x = 10$, find $\int_0^2 f(2x)\,\mathrm{d}x$.

8-23. Suppose that $f(1) = 2, f(4) = 7, f'(1) = 5, f'(4) = 3$, and $f''$ is integrable. Find $\int_1^4 xf''(x)\,\mathrm{d}x$.

8-24. Consider the graph of $y = f(x)$ on the interval $[0, 6]$ given below.



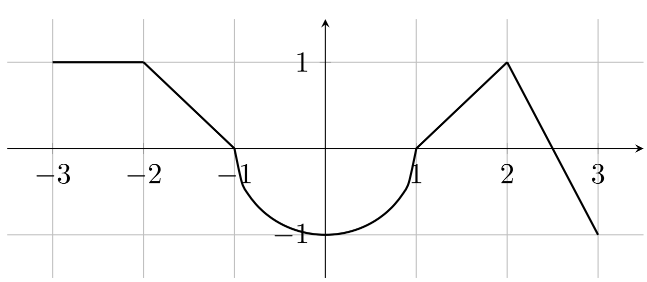

Define $G(x) = \int_1^x f(s)\,\mathrm{d}s$ on $[0, 6]$.

(a) Determine the intervals on which $G'(x) \geq 0$.

(b) Determine the inflection points of $G$.

(c) Find the absolute maximum and minimum of $G$.

8-25. Consider the graph of $y = f(x)$ on the interval $[-3, 3]$ below.



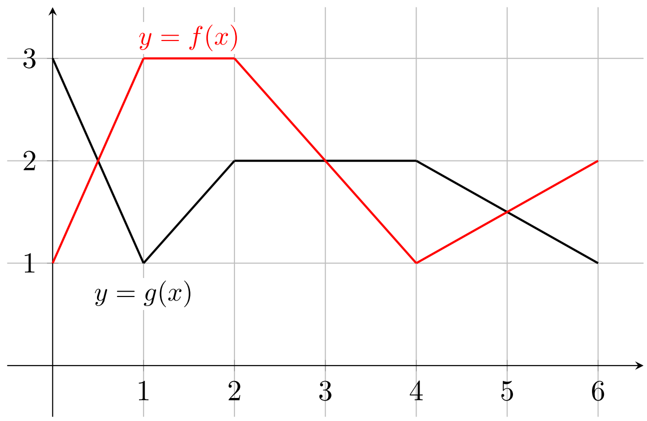
Let $H(x) = \int_0^x f(s)\,\mathrm{d}s$ on $[-3, 3]$

(a) Determine the intervals on which $H$ is increasing and decreasing.

(b) Determine the critical points of $H$.

(c) Determine the absolute maximum and minimum of $H$.

8-26. Let $f$ and $g$ be the functions plotted below on the interval $[0, 6]$.




Suppose we define the function $A(x) = \int_0^x [f(t) - g(t)]\,dt$.

(a) Determine all values of $x$ for which $A(x) > 0$.

(b) Determine all values of $x$ for which $A(x)$ is increasing/decreasing.

(c) Find the absolute maximum and absolute minimum of the function $A(x)$ on $[0, 6]$.

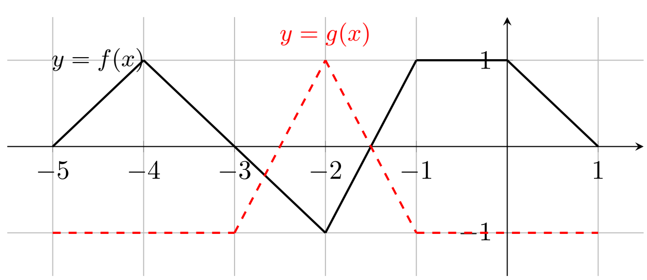
8-27. Let $f(x) = \int_x^4 (t^2 + 1)^{2020}\,\mathrm{d}t$. Compute the following:

(a) $\int_0^4 x f'(x)\,\mathrm{d}x$

(b) $\int_0^4 f(x)\,\mathrm{d}x$

(c) $\int_0^2 x^5 f(x^2)\,\mathrm{d}x$

8-28. Given the graphs of $f$ and $g$ in the figure below, evaluate the following integrals:

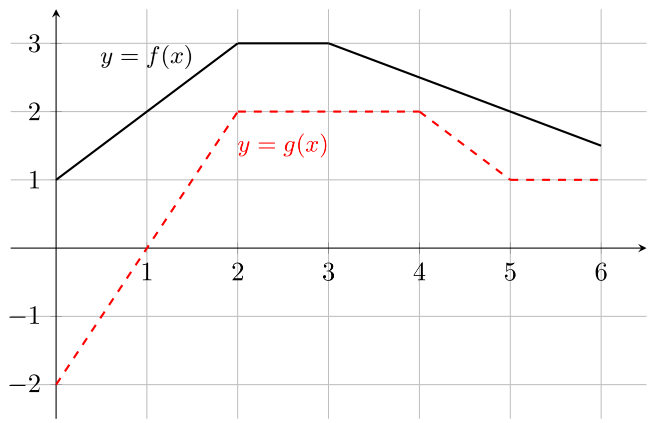


(a) $\displaystyle\int_{-2}^{0} [f(x) - g(x)]\,\mathrm{d}x$

(b) $\displaystyle\int_{-3}^{-2} [f'(x) - g'(x)]\,\mathrm{d}x$

(c) $\displaystyle\int_{-5}^{-1} f'(x)\sqrt{f(x)+1}\,\mathrm{d}x$

8-29. Let $F(x) = \displaystyle\int_0^x f(t)\,\mathrm{d}t$ and $G(x) = \displaystyle\int_0^x g(t)\,\mathrm{d}t$, where $f$ and $g$ are given in the figure. Compute:




(a) $\displaystyle\int_2^3 F(x)f(x)\,\mathrm{d}x$

(b) $\displaystyle\int_4^5 g(f(x))f'(x)\,\mathrm{d}x$

(c) $\displaystyle\int_0^5 F(x)g(x)\,\mathrm{d}x + \int_0^5 G(x)f(x)\,\mathrm{d}x$

# 9 Applications of Integration

The ability to continuously sum quantities means that integration appears in all sorts of diverse fields. Here we'll look at just a few applications of the integral.

## 9.1 Area Computations

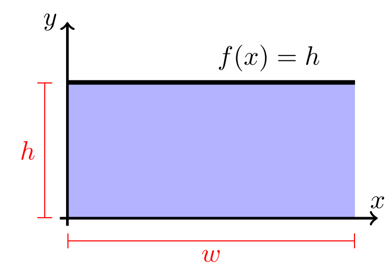
One of the primary motivations for developing the theory of integral calculus is to compute areas. The area of some objects are easy to compute, such as rectangles, triangles, parallelograms, and even trapezoids. Our ability to find formulas for the area of these shapes hinges upon the fact that they are constructed with straight lines, and so may be related to rectangles.

A quick glance around any room, let alone the free expanse of nature, very quickly confirms that there are very few naturally occurring rectangles, triangles, or trapezoids. Mother Nature, it seems, is not a fan of straight lines. You might cry, "But we know the area of a circle is $\pi r^2$!" Ah, but this formula was actually determined by Archimedes, effectively emulating integration.

### 9.1.1 What We Already Know

Let us recall that integration does give us our classical formulas:

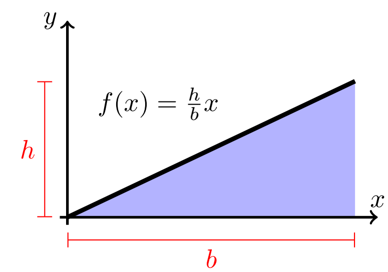
**Rectangles:**




Consider a rectangle with height $h$ and width $w$. Let $f : [0, w] \to \mathbb{R}$ be the constant function $f(x) = h$. The area under the graph of $f$ is precisely the area of the rectangle, and indeed we have

$$\int_0^w f(x)\,\mathrm{d}x = \int_0^w h\,\mathrm{d}x = hw$$

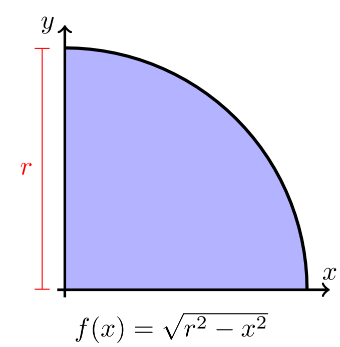
**Triangles:**




Consider a triangle with base $b$ and height $h$. Define the function $f : [0, b] \to \mathbb{R}$ by $f(x) = hx/b$, which is a straight line with height $f(b) = h$. The area under $f$ is the area of the desired rectangle, and integrating yields

$$\int_0^b \left[\frac{h}{b}x\right] \mathrm{d}x = \frac{h}{b}\frac{x^2}{2}\bigg|_0^b = \frac{1}{2}bh.$$

**Circles:**




Let $r > 0$ be the radius of our circle. We know that the formula of a circle is given by $x^2 + y^2 = r^2$. We cannot write the circle as a function though, but by writing $y = \sqrt{r^2 - x^2}$ and integrating on $[0, r]$ we can determine a quarter of the area of the circle. If we multiply by 4 at the end we will get the full area of the circle.

We don't have the ability to solve this integral, but we can nonetheless conclude that

$$\int_0^r \sqrt{r^2 - x^2}\,\mathrm{d}x = \frac{\pi r^2}{4}.$$

### 9.1.2 More Complicated Shapes

Most other shapes require that we use integration to determine their area. The next few examples are straightforward applications of the things we have learned thus far, but we really emphasize that without integration, the corresponding areas would be impossible to compute.

**Example 9.1**

Determine the area under the graph of $f(x) = xe^{x/4} + 1$ on $[0, 3]$.

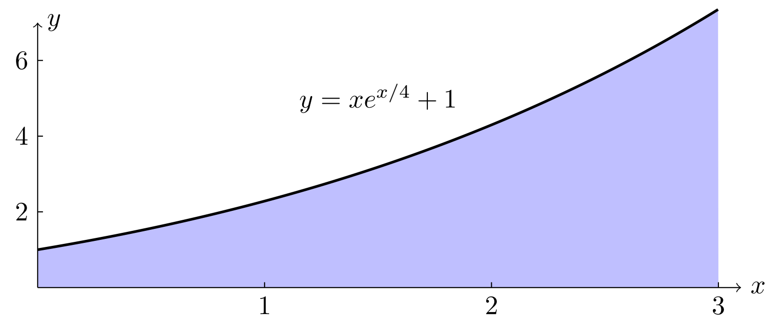



Figure 9.1: A relatively simple shape whose area is difficult to compute.

*Solution.* The plot of the corresponding area is given in Figure 9.1. Using integration by parts ($u = x$, $\mathrm{d}v = e^{x/4}\,\mathrm{d}x$), one finds that

$$\begin{aligned}\int_0^3 \left[xe^{x/4} + 1\right]\,\mathrm{d}x &= \left[4xe^{x/4} + x\right]_0^3 - 4\int_0^3 e^{x/4}\,\mathrm{d}x \\ &= \left[12e^{3/4} + 3\right] - 16\left[e^{x/4}\right]_0^3 \\ &= 19 - 4e^{3/4}.\end{aligned}$$

■

**Example 9.2**

Determine the area under the graph of $f(x) = x\sqrt{1-x^2}$ on $[0,1]$.

*Solution.* Evaluating the integral, we get

$$\begin{aligned}\int_0^1 x\sqrt{1-x^2}\,\mathrm{d}x &= -\frac{1}{2}\int_1^0 \sqrt{u}\,\mathrm{d}u & u &= 1-x^2 \\ & & \mathrm{d}u &= -2x\,\mathrm{d}x \\ &= \left[\frac{1}{3}u^{3/2}\right]_0^1 = \frac{1}{3}.\end{aligned}$$

■

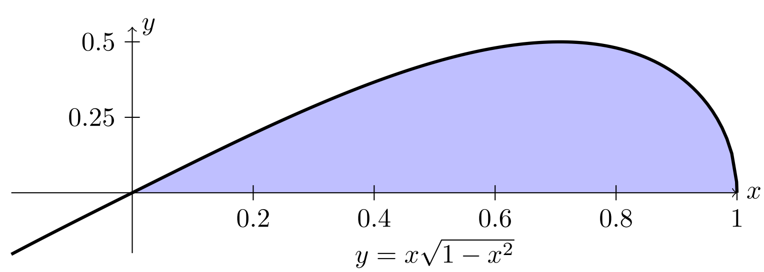



Figure 9.2: Plot of the function $f(x) = x\sqrt{1-x^2}$ for Example 9.2.

### 9.1.3 Unsigned (Absolute) Area

By now we are quite familiar with the fact that integrals compute *signed* areas; that is, areas above the $x$-axis are given positive sign, while those beneath the $x$-axis carry a negative sign.

To compute the unsigned, absolute area, we modify our function so that formerly negative areas now lie above the $x$-axis. This is done by taking the absolute value of the integrand. If $f : [a,b] \to \mathbb{R}$ is integrable, then

$$\text{Absolute Area} = \int_a^b |f(x)| \, \mathrm{d}x.$$

**Example 9.3**

Determine the unsigned (that is, total) area under the graph $y = x-2$ on the interval $[0,10]$.

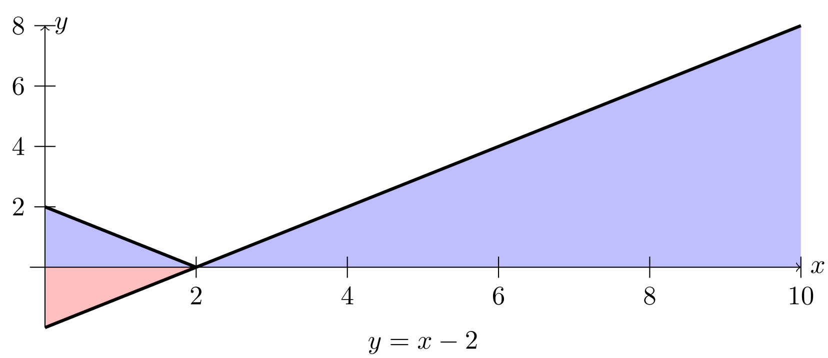



Figure 9.3: The function $f(x) = x - 2$ is negative on $[0,2]$ and positive on $[2,10]$. The absolute area under the graph must account for the signed difference.

*Solution.* The absolute area is determined by integrating the absolute value $f(x) = |x-2|$. This is equivalently $2-x$ on $[0,2]$ and $x-2$ on $[2,10]$, so

$$\int_0^{10} |x-2| \, \mathrm{d}x = \int_0^2 |x-2| \, \mathrm{d}x + \int_2^{10} |x-2| \, \mathrm{d}x = \int_0^2 2-x \, \mathrm{d}x + \int_2^{10} x-2 \, \mathrm{d}x = 2 + 32 = 34.$$

The picture corresponding to this problem is shown in Figure 9.3. ■

**Example 9.4**

Determine the total area beneath the graph of the function $f(x) = x^2 - 1$ on the interval $[-2,2]$.

*Solution.* We may determine the absolute area by integrating $|f(x)| = |x^2 - 1|$, but in practice this requires that we determine where $f(x) < 0$. You can quickly verify that the roots of $f$ lie at

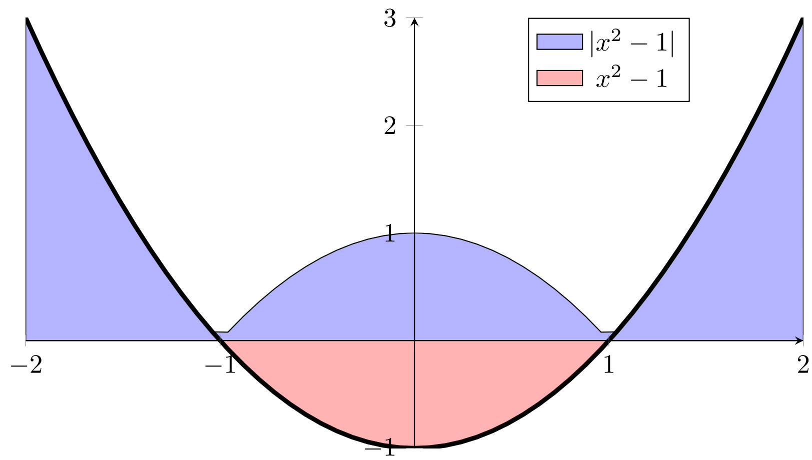



Figure 9.4: The blue area is that which is integrated by taking the absolute value. The red area on $[-1,1]$ is the same as the blue area directly above it, except the blue carries a sign of $+1$ while the red carries a sign of $-1$.

$x = \pm 1$, and that $f(x) < 0$ on $[-1,1]$. Furthermore, $|f|$ is an even function, allowing us to perform the integral on just $[0,2]$:

$$\begin{aligned}\int_{-2}^{2} |x^2-1|\,\mathrm{d}x &= 2\int_0^2 |x^2-1|\,\mathrm{d}x = 2\left[\int_0^1 \left[1-x^2\right]\,\mathrm{d}x + \int_1^2 \left[x^2-1\right]\,\mathrm{d}x\right] \\ &= 2\left[x - \frac{x^3}{3}\right]_0^1 + 2\left[\frac{x^3}{3} - x\right]_1^2 = 4.\end{aligned}$$

■

### 9.1.4 Integrating along the $y$-axis

There may be occasions where one is interested in the area under a curve which cannot necessarily be represented by a function, or in which it is merely inconvenient to write as a function. Examples like this will manifest in Section 9.1.5. As such, it may be more useful to integrate along the $y$-axis rather than the $x$-axis.

For example, consider the curve described by $y^2 - x - 2 = 0$, which is plotted in Figure 9.5. The curve does not describe a function in $x$, though we can solve for the two function components $y = \pm\sqrt{x+2}$. On the other hand, the curve is a function in $y$ as we can write $x = y^2 - 2$. Were we set up a Riemann sum for this function, it would look like Figure-9.5. Notice that this is not the same area one would find if we were to integrate with respect to $x$ (Figure-9.6).

**Example 9.5**

Compare the areas under the curve $y^2 - x - 4 = 0$ when integrated along the $x$- and $y$-direction, on the interval $x \in [-4,5]$.

*Solution.* We first integrate along the $x$-direction as usual. While the curve is not described by a

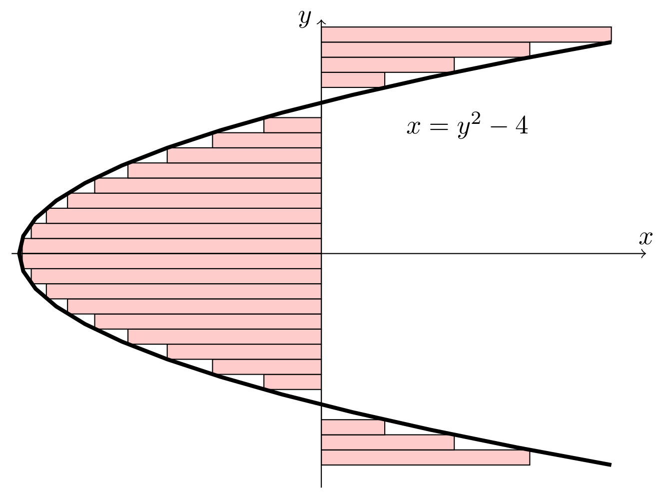



Figure 9.5: A Riemann sum for $x = y^2 - 4$, treated as a function of $y$.

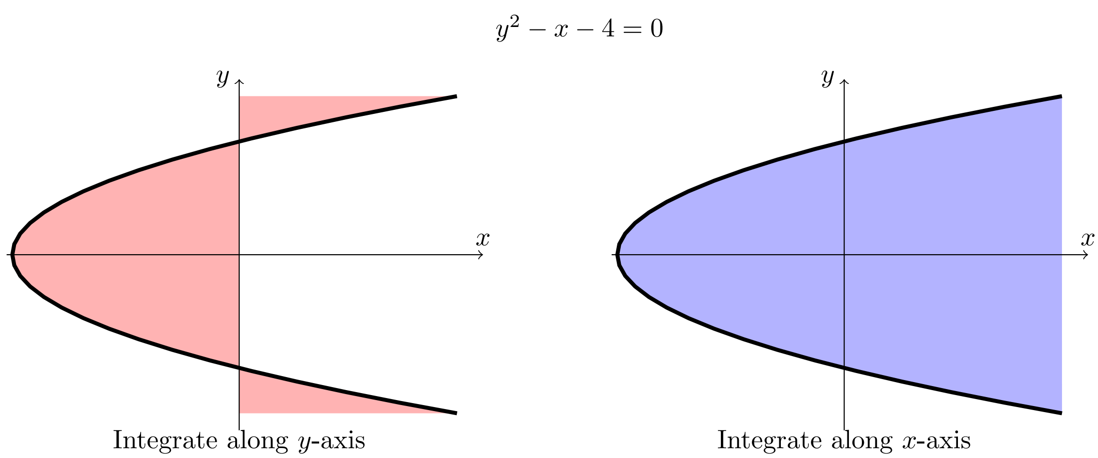



Figure 9.6: The computed area depends upon whether we are integrating with respect to $x$ or $y$.

function, we can just integrate $y = \sqrt{x+4}$ and double the final answer.

$$2\int_{-4}^{5} \sqrt{x+4}\,\mathrm{d}x = 3\,(x+4)^{3/2}\Big|_{-4}^{5} = 81.$$

On the other hand, if we set $x = y^2 - 4$ then $x \in [-4, 5]$ implies that $y \in [-3, 3]$, so

$$\int_{-3}^{3} \left(y^2 - 4\right)\,\mathrm{d}y = 2\int_{0}^{3} \left(y^2 - 4\right)\,\mathrm{d}y = 2\left[\frac{y^3}{3} - 4y\right]_0^3 = -6.$$

A very significant difference. ■

### 9.1.5 The Area Between Curves

Every example thus far measured the area between the graph of a function and the $x$-axis. We can increase our flexibility with a bit of creativity. For example, let's say we want to find the area bounded by the curves $y = x^2$ and $y^2 = x$ (Figure 9.7).

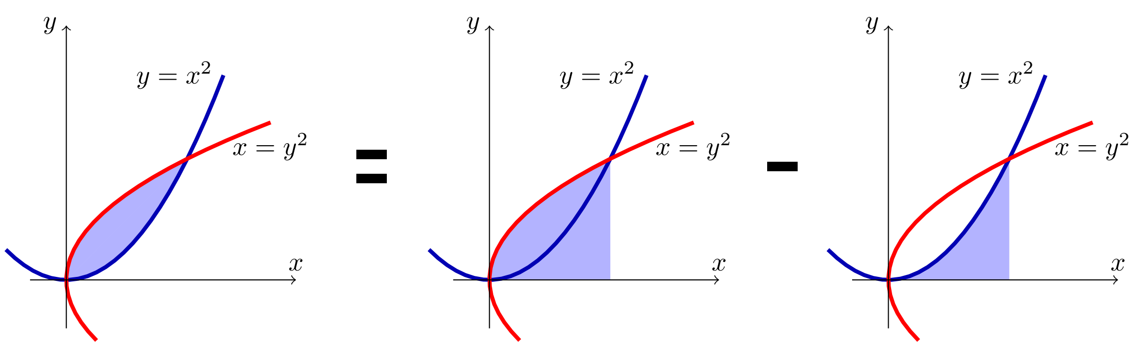



Figure 9.7: To compute the area bounded between the curves $y = x^2$ and $x = y^2$ requires that we be more creative.

The idea is fairly simple: By computing the area under $y = x^2$ and subtracting the area under $x = y^2$, we should get the area bounded between the two curves. More generally, given two curves $f, g$ on $[a, b]$, the area between $f$ and $g$ can be computed as

$$\int_a^b [f(x) - g(x)]\,\mathrm{d}x. \tag{9.1}$$

We note though that this is still a *signed area*. In particular, any interval where $f(x) > g(x)$ will be assigned a positive area, while area where $f < g$ will be given a negative area. Of course, the total area can be computed by $\int_a^b |f(x) - g(x)|\,\mathrm{d}x$.

Equation (9.1) can also be interpreted as the limit of a Riemann sum, or as integrating the function $f - g$. In the case where $f(x) = \sqrt{x}$ and $g(x) = x^2$, these two pictures are given by Figure 9.8.

**Example 9.6**

Find the area bounded between the functions $f(x) = \sqrt{x}$ and $g(x) = x^2$.

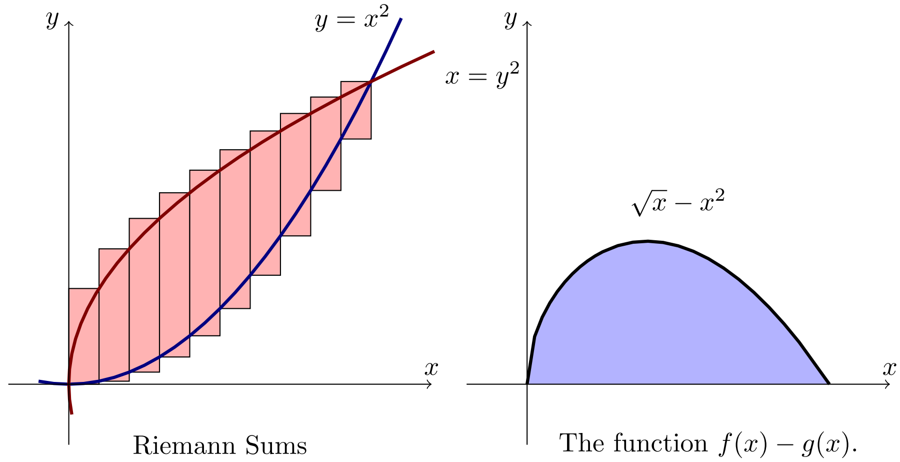



Figure 9.8: Two different interpretations of $\int_a^b [f(x) - g(x)]\, dx$. Left: As a Riemann sum. Right: As the integral of the function $f - g$.

*Solution.* Notice that we were not explicitly given an interval over which to integrate. The reason is that if one sketches the graphs of $f$ and $g$, there is only one area that can be said to be enclosed by the two functions. Consequently, we need to determine where the appropriate intersections occur. This can be done by equation $f(x) = g(x)$.

Setting $x^2 = \sqrt{x}$, one can easily to solve to find that the intercepts occur at $x = 0$ and $x = 1$. Furthermore, on $[0, 1]$ we have that $\sqrt{x} > x^2$, and hence our area is given as

$$\int_0^1 \left[\sqrt{x} - x^2\right] dx = \left[\frac{2}{3}x^{3/2} - \frac{1}{3}x^3\right]_0^1 = \frac{2}{3} - \frac{1}{3} = \frac{1}{3}.$$

We can also integrate in the $y$-direction. Here our interval is still $[0, 1]$ but $y^2 < \sqrt{y}$, so

$$\int_0^1 \left[\sqrt{y} - y^2\right] dy = \frac{1}{3}$$

gives exactly the same answer. ■

**Example 9.7**

Determine the area bounded by the curves $y^2 = 2x + 6$ and $y = x - 1$.

*Solution.* To integrate with respect to $x$ we would need to break the curves down into the regions where the difference can be written as a function. This is doable, but is fairly complicated. Instead, if we integrate with respect to $y$ this becomes rather simple.

We must first determine where the two curves intersect. Substituting $y = x - 1$ into $y^2 = 2x + 6$ we get

$$(x - 1)^2 = 2x + 6 \quad \Leftrightarrow \quad x^2 - 4x - 5 = 0 \quad \Leftrightarrow \quad x = 5, -1.$$

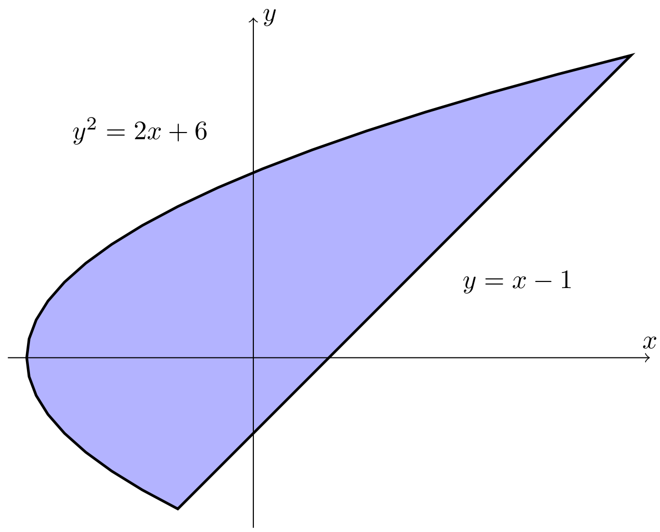



Figure 9.9: The area between the curves $y^2 = 2x + 6$ and $y = x - 1$.

Computing the corresponding $y$-values, we get that the intercept points occur at $(-1, -2)$ and $(5, 4)$. Hence $y$ ranges over the interval $[-2, 4]$. Rewriting our curves as functions of $y$, we get $x = y + 1$ and $x = y^2/3 - 3$. The area can be computed to be

$$\begin{aligned}\int_{-2}^{4}\left[(y+1) - \left(\frac{y^2}{3} - 3\right)\right] \mathrm{d}y &= \int_{-2}^{4}\left[-\frac{y^2}{3} + y + 4\right] \mathrm{d}y = \left[-\frac{y^3}{9} + \frac{y^2}{2} + 4y\right]_{-2}^{4} \\ &= \left(\frac{64}{9} + 8 + 16\right) - \left(\frac{8}{9} + 2 - 9\right) = 22.\end{aligned}$$ ■

## 9.2 Improper Integrals

It was essential in defining the integral that we examined *bounded* functions on *bounded* intervals $[a, b]$, so that everything under consideration could be finite. In this section we examine how to now extend the idea of integrals to cover unbounded functions, and functions defined on unbounded intervals.

### 9.2.1 Infinite Intervals

Our goal is to extend the notion of integration from a finite interval $[a, b]$ to an infinite interval $[a, \infty)$ or $(-\infty, b]$. We will develop the idea for the interval $[a, \infty)$ and leave the details for $(-\infty, b]$ as an exercise.

**Definition 9.8**

Let $f$ be a bounded function on the interval $[a, \infty)$ such that $f$ is integrable on $[a, x]$ for every $x > a$. We define the *improper integral* of $f$ on $[a, \infty)$ as

$$\int_a^\infty f(t)\,\mathrm{d}t = \lim_{x\to\infty} \int_a^x f(t)\,\mathrm{d}t.$$

We say that that the improper integral *converges* if this limit is finite, and *diverges* otherwise.

Let us take a moment to think about what this is saying: We are defining a new function

$$F(x) = \int_a^x f(t)\,\mathrm{d}t,$$

which is an anti-derivative of $f$. The improper integral then converges if the function $F$ has a horizontal asymptote; that is, the area under the graph of $f$ asymptotically stabilizes to a single, finite number.

Naturally, one then defines the improper integral on $(-\infty, b]$ as

$$\int_{-\infty}^b f(t)\,\mathrm{d}t = \lim_{x\to-\infty} \int_x^b f(t)\,\mathrm{d}t.$$

**Example 9.9**

Determine $\displaystyle\int_0^\infty e^{-t}\,\mathrm{d}t$, if it exists.

*Solution.* By definition, we know that $\displaystyle\int_0^\infty e^{-t}\,\mathrm{d}t = \lim_{x\to\infty}\int_0^x e^{-t}\,\mathrm{d}t$. We are familiar with computing the integral on the right hand side, and know that $\int_0^x e^{-t}\,\mathrm{d}t = -e^{-t}\big|_0^x = 1 - e^{-x}$. Thus

$$\int_0^\infty e^{-t}\,\mathrm{d}t = \lim_{x\to\infty}\int_0^x e^{-t}\,\mathrm{d}t = \lim_{x\to\infty}\left[1 - e^{-t}\right] = 1. \qquad \blacksquare$$

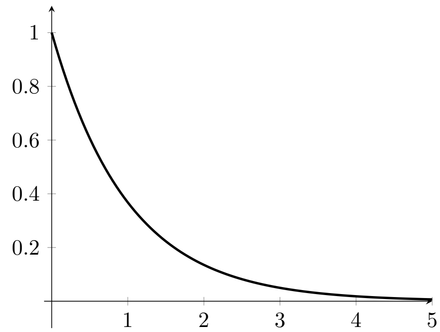
**Example 9.10**

Determine $\displaystyle\int_0^\infty \frac{1}{\sqrt{t}}\,\mathrm{d}t$, if it exists.

*Solution.* Proceeding by definition, we have that

$$\int_0^\infty \frac{1}{\sqrt{t}}\,\mathrm{d}t = \lim_{x\to\infty}\int_0^x \frac{1}{\sqrt{t}}\,\mathrm{d}t = \lim_{x\to\infty}\left[2\sqrt{x}\right],$$

which does not exist, since it diverges off to infinity. $\blacksquare$



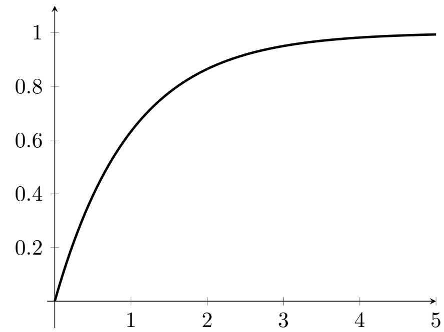




Figure 9.10: The function $f(x) = e^{-x}$ and an anti-derivative $F(x) = 1 - e^{-x}$. We see that $F(x)$ tends to the number 1 as $x \to \infty$. This occurs because as $x \to \infty$ the graph under the function $f(x)$ becomes very small.

Intuitively, it seems like functions which tend to zero should have a convergent improper integral, but Example 9.10 shows this need not be the case. Is there a line at which rational functions $x^p$ transition from convergent to divergent?

**Proposition 9.11**

If $a > 0$ is an arbitrary positive number, then

$$\int_a^\infty \frac{1}{x^p}\,\mathrm{d}x \qquad \text{converges if and only if} \qquad p > 1.$$

*Proof.* If $p = 1$ then

$$\int_a^\infty \frac{1}{t}\,\mathrm{d}t = \lim_{x\to\infty} \ln(x/a) = \infty$$

so the integral diverges. Thus assume that $p \neq 1$, for which we have

$$\int_a^\infty \frac{1}{t^p}\,\mathrm{d}t = \lim_{x\to\infty} \left[\frac{1}{1-p}\frac{1}{x^{p-1}}\right]_a^x.$$

We know that $1/x^{p-1}$ converges only if and if the power is non-negative; that is, $p - 1 \geq 0$. Combining this with $p \neq 1$ tells us that the improper integral converges if and only if $p > 1$. □

As was computed explicitly, this implies that $\int_1^\infty 1/x\,\mathrm{d}x$ does not converge, which is a result that often confuses students.

What happens if we want to define the improper integral on $(-\infty, \infty)$?

**Definition 9.12**

If $f$ is integrable on every interval $[a,b] \subseteq \mathbb{R}$, then we say that $\displaystyle\int_{-\infty}^{\infty} f(t)\,\mathrm{d}t$ converges if, for any $c \in \mathbb{R}$ we have both

$$\int_{-\infty}^{c} f(t)\,\mathrm{d}t \text{ converges, and } \int_{c}^{\infty} f(t)\,\mathrm{d}t \text{ converges.}$$

In this case, we set[a]

$$\int_{-\infty}^{\infty} f(t)\,\mathrm{d}t = \int_{-\infty}^{c} f(t)\,\mathrm{d}t + \int_{c}^{\infty} f(t)\,\mathrm{d}t.$$

[a] The student should convince him/herself that the value of the improper integral does not depend on the value of $c$.

This is different than simply demanding that

$$\int_{-\infty}^{\infty} f(t)\,\mathrm{d}t = \lim_{x\to\infty}\int_{-x}^{x} f(t)\,\mathrm{d}t \qquad \text{exists.}$$

The reason is that the value of the integral should not depend on 'how quickly' we take our limits. For example

$$\lim_{x\to\infty}\int_{-x}^{x} t\,\mathrm{d}t = \lim_{x\to\infty}\left[\frac{1}{2}x^2 - \frac{1}{2}x^2\right] = 0.$$

On the other hand,

$$\lim_{x\to\infty}\int_{-x}^{2x} t\,\mathrm{d}t = \lim_{x\to\infty}\left[2x^2 - \frac{1}{2}x^2\right] = \lim_{x\to\infty}\frac{3}{2}x^2 \qquad \text{does not exist} \quad .$$

For the limit to make sense, these two quantities should be the same and they are clearly not.

**Example 9.13**

Determine $\displaystyle\int_{-\infty}^{\infty} e^{-|t|}\,\mathrm{d}t$.

*Solution.* A natural place to split our interval will be at 0. Now

$$\begin{aligned} \int_{0}^{\infty} e^{-|t|}\,\mathrm{d}t &= \int_{0}^{\infty} e^{-t}\,\mathrm{d}t && \text{since } t > 0 \\ &= 1 && \text{by Example 9.9.} \end{aligned}$$

Similarly, $\displaystyle\int_{-\infty}^{0} e^{-|t|}\,\mathrm{d}t = 1$, thus

$$\int_{-\infty}^{\infty} e^{-|t|}\,\mathrm{d}t = \int_{-\infty}^{0} e^{-|t|}\,\mathrm{d}t + \int_{0}^{\infty} e^{-|t|}\,\mathrm{d}t = 2. \qquad \blacksquare$$

### 9.2.2 Unbounded Functions

The case of unbounded functions often poses even more difficulty, since it is very tempting to just blindly apply the Fundamental Theorem of Calculus without paying attention.

**Example 9.14**

Compute the integral $\displaystyle\int_{-1}^{1} \frac{1}{x^2}\,\mathrm{d}x$.

*Solution.* We know that an anti-derivative of $1/x^2$ is $-1/x$, so if we were to just blindly apply the FTC we would get

$$\int_{-1}^{1} \frac{1}{x^2}\,\mathrm{d}x = -\frac{1}{x}\Big|_{-1}^{1} = -2.$$

Unfortunately, **this is completely and totally wrong.** Our first hint at a miscalculation is probably the fact that $1/x^2$ is everywhere positive, yet we somehow ended up with a negative integral. In fact, the integral is infinite. To see this, note that since $\frac{1}{x^2} > 0$ then

$$\int_{-1}^{1} \frac{1}{x^2}\,\mathrm{d}x \geq \int_{\epsilon}^{1} \frac{1}{x^2}\,\mathrm{d}x = -\frac{1}{x}\Big|_{\epsilon}^{1} = \frac{1}{\epsilon} - 1,$$

and that by choosing $\epsilon$ to be small enough we can make the integral arbitrarily large. The reason is that the function $1/x^2$ is not integrable on the interval $[-1,1]$ – all integrable functions are bounded – and hence we could not apply the FTC. ■

**Exercise:** Compare the following two expressions:

$$\lim_{\epsilon\to 0^+}\left[\int_{-1}^{-\epsilon} \frac{1}{x}\,\mathrm{d}x + \int_{\epsilon}^{1} \frac{1}{x}\,\mathrm{d}x\right] \qquad \lim_{\epsilon\to 0^+}\left[\int_{-1}^{-\epsilon} \frac{1}{x}\,\mathrm{d}x + \int_{2\epsilon}^{1} \frac{1}{x}\,\mathrm{d}x\right]$$

The way to deal with unbounded functions is precisely the same way that we deal with unbounded intervals: we take a limit.

**Definition 9.15**

If $f$ is a function on $[a,b]$, unbounded at $a$, but integrable on $[x,b]$ for every $x > a$ then we define the improper integral

$$\int_a^b f(t)\,\mathrm{d}t = \lim_{x\to a^+} \int_x^b f(t)\,\mathrm{d}t.$$

If this limit is finite we say that the improper integral *converges*; otherwise, we say that the improper integral *diverges*.

Similarly, if $f$ were unbounded at $b$, we would define the improper integral as

$$\int_a^b f(t)\,\mathrm{d}t = \lim_{x\to b^-} \int_a^x f(t)\,\mathrm{d}t.$$

**Example 9.16**

Evaluate $\displaystyle\int_3^5 \frac{t}{\sqrt{t^2-9}}\, dt$, if it exists.

*Solution.* We immediately recognize that there could be a problem at $t = 3$. Using the substitution $u = t^2 - 9$ we get

$$\int_3^5 \frac{t}{\sqrt{t^2-9}}\, dt = \lim_{x\to 3^+} \left[\sqrt{t^2-9}\right]_x^5 = \lim_{x\to 3^+} \left[4 - \sqrt{x^2-9}\right] = 4. \qquad \blacksquare$$

Just as in the case of integrating from $(-\infty, \infty)$ we must also be careful about integrating on both sides of an unbounded function.

**Definition 9.17**

If $f$ is a function on $[a, b]$ which is unbounded at $c \in [a, b]$ then we say that the improper integral

$$\int_a^b f(t)\, dt \quad \text{converges, if and only if} \quad \int_a^c f(t)\, dt \quad \text{and} \quad \int_c^b f(t)\, dt \quad \text{both exist.}$$

In this case, we set

$$\int_a^b f(t)\, dt = \int_a^c f(t)\, dt + \int_c^b f(t)\, dt.$$

**Proposition 9.18**

For any $a \neq 0$ we have that

$$\int_0^a \frac{1}{t^p}\, dt \text{ converges, if and only if } p < 1.$$

*Proof.* For simplicity, let's assume that $a > 0$. If $p = 1$ then

$$\int_0^a \frac{1}{t}\, dt = \lim_{x\to 0^+} \int_x^a \frac{1}{t}\, dt = \lim_{x\to 0^+} \ln(t)\Big|_x^a = \infty.$$

If $p \neq 1$ then

$$\int_0^1 \frac{1}{t^p}\, dt = \lim_{x\to 0^+} \int_x^a \frac{1}{t}\, dt = \frac{1}{1-p} \lim_{x\to 0^+} \frac{1}{x^{p-1}}\Big|_x^a.$$

The limit converges if and only if $p - 1 \leq 0$ so that $p \leq 1$. Combined with the fact that $p \neq 1$ we get $p < 1$ as required. $\square$

### 9.2.3 The Basic Comparison Test

In this section, we develop some techniques to make our lives simpler in terms of dealing with improper integrals. The idea is something like the following: Say that you were asked to determine whether

$$\int_1^\infty \frac{x^2+1}{x^4+3x^2-4x+1}\,\mathrm{d}x$$

converges. This integral is not easy to compute explicitly. The goal, as is often the case with mathematics, is to reformulate such a problem into one that is easier to solve, or that has already been solved. The Basic Comparison Test is the simplest of the basic tests, and exploits the "monotonicity" of the integral.

**Theorem 9.19: The Basic Comparison Test for Improper Integrals**

Let $f, g$ be functions on an interval $[a,\infty)$ such that $0 \leq f(x) \leq g(x)$ for all $x \in [a,\infty)$.

1. If $\displaystyle\int_a^\infty g(t)\,\mathrm{d}t$ converges then $\displaystyle\int_a^\infty f(t)\,\mathrm{d}t$ converges.
2. If $\displaystyle\int_a^\infty f(t)\,\mathrm{d}t$ diverges then $\displaystyle\int_a^\infty g(t)\,\mathrm{d}t$ diverges.

The idea is again a type of 'Squeeze Theorem' argument. If the integral of the bigger function $g$ becomes finite, the monotonicity of the integral cannot allow $f$ to go off to infinity. Similarly, if the integral of the smaller function $f$ goes off to infinity, the larger function's integral must also diverge. A good question at this point is to ask whether any of the integrals could oscillate and hence not converge. The condition that $0 \leq f(x) \leq g(x)$ guarantees that the integrands are positive, and hence that the corresponding integrals are increasing functions.

**Example 9.20**

Show that $\displaystyle\int_1^\infty \frac{2+e^x}{x}\,\mathrm{d}x$ diverges.

*Solution.* We want to compare the function $(2+e^x)/x$ to some function which we know diverges. Since $e^x > 1$ for all $x > 1$, we have

$$\frac{2+e^x}{x} \geq \frac{2}{x}.$$

By Proposition **??**, we know that $\int_1^\infty \frac{2}{x}\,\mathrm{d}x$ diverges, so by the Comparison Test it follows that

$$\int_1^\infty \frac{2+e^x}{x}\,\mathrm{d}x \text{ diverges.}$$

■

**Example 9.21**

Determine whether $\displaystyle\int_0^\infty \frac{x}{\sqrt{x^6+1}}\,\mathrm{d}x$ converges or diverges.

*Solution.* Once again, we just want to look at which terms in the numerator and denominator dominate in the limit $x \to \infty$. Certainly, we do not expect $x^6+1$ to be too much different than $x^6$

for very large $x$, so we will compare our integrand to the function

$$\frac{x}{\sqrt{x^6}} = \frac{1}{x^2}.$$

Indeed, since $1 + x^6 \geq x^6$ we have that $\frac{1}{\sqrt{x^6}} \geq \frac{1}{\sqrt{1+x^6}}$, which in turn implies that

$$\frac{1}{x^2} = \frac{1}{\sqrt{x^6}} \geq \frac{x}{\sqrt{x^6+1}}.$$

Now the integral of the left-hand-side converges by **??**, so by the Comparison Test we know that

$$\int_0^\infty \frac{x}{\sqrt{x^6+1}}\,\mathrm{d}x \text{ converges.}$$ ■

### 9.2.4 The Limit Comparison Test

The Basic Comparison Test is just that, basic. Often times the obvious inequality that you want actually ends up going in the wrong direction, yet the integrals are so similar that you feel like you should still be able to compare them. Example 9.23 below will demonstrate precisely this.

The Limit Comparison Test will fix this by asking the question: "Do $f$ and $g$ grow at roughly the same rate?"

**Theorem 9.22: The Limit Comparison Test**

Let $f, g$ be non-negative integrable functions on all subintervals of $[a, \infty)$. If

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

then $\displaystyle\int_a^\infty f(x)\,\mathrm{d}x$ converges if and only if $\displaystyle\int_a^\infty g(x)\,\mathrm{d}x$ converges.

The statement that $f/g$ converges to some finite, non-zero number, means that $f$ and $g$ grow asymptotically at the same speed, up to some multiplicative constant (which is precisely the value of the limit). Note that if the limit is 0, eventually one must have $g(x) \geq f(x)$ and can use the Basic Comparison Test appropriately. Similarly, if the limit is $\infty$, then eventually $f(x) \geq g(x)$ and again the Basic Comparison Test can be used.

Also notice that the ratio does not matter, for if $f(x)/g(x) \to L$ which is finite and positive, then $g(x)/f(x) \to \frac{1}{L}$ which is also finite and positive.

**Example 9.23**

Determine whether $\displaystyle\int_1^\infty \frac{1}{\sqrt{1+x}}\,\mathrm{d}x$ converges or diverges.

*Solution.* If one were to try to use the Basic Comparison Test, the obvious inequality is that $1/\sqrt{1+x} \leq 1\sqrt{x}$. But this does not tell us anything! The right-hand-side diverges, and so does not impose its will on the left-hand-side. Instead, we recognize that

$$\lim_{x\to\infty} \frac{\sqrt{x}}{\sqrt{1+x}} = 1.$$

Thus by the Limit Comparison Test, since $\int_1^\infty 1/\sqrt{x}\,\mathrm{d}x$ diverges, we necessarily have that $\int_1^\infty 1/\sqrt{1+x}\,\mathrm{d}x$ diverges as well. ■

**Example 9.24**

Determine whether $\displaystyle\int_1^\infty \frac{x^2+1}{x^4+3x^2-4x+1}\,\mathrm{d}x$ converges or diverges.

*Solution.* With a strong enough argument, one might be able to argue this example using the Basic Comparison Test, but the Limit Comparison Test proves much simpler. Again the idea is to look at how the numerator and denominator grow asymptotically. The numerator grows as $x^2$, while the denominator grows as $x^4$, meaning that the combined system grows as $1/x^2$. To invoke the Limit comparison Test, we must compute the limit of the ratio of these functions:

$$\lim_{x\to\infty} \frac{x^2+1}{x^4+3x^2-4x+1}\frac{1}{1/x^2} = \lim_{x\to\infty} \frac{x^4+x^2}{x^4+3x^2-4x+1} = 1.$$

Since $\int_1^\infty 1/x^2\,\mathrm{d}x$ converges (by Proposition **??**), we conclude by the Limit Comparison Test that

$$\int_1^\infty \frac{x^2+1}{x^4+3x^2-4x+1}\,\mathrm{d}x \text{ also converges.}$$ ■

## 9.3 Applications in Economics and Finance

We've seen that using an exponential model of compounding growth can simplify the theory and manipulations needed to compute desired quantities. An obstacle with this model is that it assumes continuously compounded interest. If we have an annuity which is continuously invested, how can we determine its future and present values?

Consider the problem of an annuity which is invested continuously at a rate $r\%$ such that in each year $R$ dollars is deposited. Let's determine the value of the annuity after $n$ years. To do this, we take a uniform partition of the interval $[0,n]$ into say $m$ subintervals. Each subinterval has length $\Delta t = n/m$. We approximate the continuous annuity by using a discrete annuity invested at the beginning of each period $t_k = k\Delta t$ for $k = 1,\ldots,n$. At time $t_k$, we deposit $R\Delta t$ dollars into the account, the future value of this money compounded continuously is $R\Delta t e^{r(n-t_k)}$, and the total future value of all deposits is the sum

$$\sum_{k=1}^n Re^{n-t_k}\Delta t$$

which is the right Riemann sum for the function $f(t) = Re^{r(n-t)}$ on $[0,n]$. Taking the limit as $m\to\infty$, we thus get

$$\begin{array}{r}\text{Future Value of Continuously}\\ \text{Invested Annuity}\end{array} = \int_0^n Re^{r(n-t)}\,\mathrm{d}t.$$

**Example 9.25**

If an annuity is invested continuously at a rate of 6% such that 20000 is invested each year, determine the value of the annuity after 3 years.

*Solution.* Using the example above, the future value of the annuity is

$$\int_0^3 20000e^{0.06(3-t)}\,\mathrm{d}t = -\frac{20000}{0.06}e^{0.06(3-t)}\Big|_{t=0}^{3} = -\frac{20000}{0.06}[1-e^{0.18}] \approx 65739.10. \qquad \blacksquare$$

Suppose you have two investment plans which generate profit $P_1(t)$ and $P_2(t)$ as functions of time $t$, with rates of profitability $P_1'$ and $P_2'$. The excess profit of one plan over another at a single instant in time is $E(t) = P_2(t) - P_1(t)$. The *net excess profit* over the time period $0 \leq t \leq n$ is $E(n) - E(0)$, or equivalently

$$E(n) - E(0) = \int_0^n E'(t)\,\mathrm{d}t = \int_0^n \left[P_2'(t) - P_1'(t)\right]\,\mathrm{d}t.$$

**Example 9.26**

Consider two investment plans whose rate of profitability is governed by

$$r_1(t) = 25 + 2t^2 \quad \text{and} \quad r_2(t) = 125 + 10t,$$

where $t$ is given in years.

1. How long does the rate of profitability of $r_1$ exceed $r_2$?
2. Determine the net excess profit during the time interval found in part (a).

*Solution.*

1. We begin by solving $r_2(t) = r_1(t)$, which when we equate corresponds to

$$25 + 2t^2 = 125 + 10t \quad \Rightarrow \quad 2t^2 - 10t - 100 = 2(t-10)(t+5) = 0.$$

The zeros occur at $t = -5$ and $t = 10$, but $t = -5$ is nonsense and so is discarded. Thus Plan 2 is more profitable for 10 years.

2. The net excess profit is

$$\begin{aligned}\int_0^{10} [125+10t] - [25+2t^2]\,\mathrm{d}t &= \int_0^{10} 100 + 10t - 2t^2\,\mathrm{d}t = \left[100t + 5t^2 - \frac{2}{3}t^3\right]_0^{10} \\ &= \frac{2500}{3} = 833.33.\end{aligned} \qquad \blacksquare$$

The consumer demand function $D$ describes the price *per unit* $p$ that a consumer is willing to spend to by $q$ units, with the implicit understanding that this is a decreasing function. For example, a consumer might buy 1 laptop for \$1000, but 2 laptops for \$1500. In the former case, the per-quantity price of the laptop is \$1000 while in the latter it is \$750.

Suppose then that we know the consumer demand function $D$. Of more value might be the *total willingness to spend* function $A(q)$, which describes the total amount a consumer is willing to spend to buy $q$ units. In this sense, the consumer demand function is the marginal willingness to spend; that is,

$$\frac{\mathrm{d}A}{\mathrm{d}q} = D(q).$$

If we know $D$, we can therefore find $A$ buy integrating.

**Example 9.27**

It is known that the consumer demand function for buying televisions is $D(q) = 200(5 - q^2)$. Find the total amount of money consumers are willing to spend to buy 3 televisions.

*Solution.* Using the fact that the consumer demand function is the marginal willingness to spend, we get

$$\int_0^3 D(q)\,\mathrm{d}q = \int_0^3 200(5-q^2)\,\mathrm{d}q = 200\left[5q - \frac{q^3}{3}\right]_0^3 = 1200.$$

Hence consumers are willing to spend \$1200 to buy three televisions. ∎

Consumer surplus measures consumer savings relative to market prices. For example, if a consumer is will to spend \$1200 to buy three televisions, but market prices dictate that you can buy three televisions for \$1000, then you have a \$200 surplus. Once again, we use the consumer demand curve $p = D(q)$. If the market dictates a price $p_0$, the total the consumer spends is $p_0 q_0$ where $q_0$ is the corresponding quantity the consumer buys. On the other hand, the consumer would have spent $A(q)$, so the consumer surplus is

$$A(q) - p_0 q_0 = \int_0^{q_0} \left[D(q) - p_0\right]\,\mathrm{d}q.$$

Producers surplus is the same concept but in the other direction. Suppose $S$ is a supply function, so that $p = S(q)$ describes the price per unit that a producer is willing to accept to supply $q$ units. The producers surplus is the difference between what they are willing to accept for supplying $q_0$ units versus what they receive in reality. At a price point of $p_0$, a producer actually receives $p_0 q_0$ dollars, while they would have been willing to accept $\int_0^{q_0} S(q)\,\mathrm{d}q$. Hence the producer surplus is

$$p_0 q_0 - \int_0^{q_0} S(q)\,\mathrm{d}q = \int_0^{q_0} \left[p_0 - S(q)\right]\,\mathrm{d}q.$$

Note that the market prices are indicated by the equilibrium point, where $D(q) = S(q)$ (that is, supply meets demand).

**Example 9.28**

Suppose the supply and demand functions are given by

$$S(q) = \frac{q^2}{3} + 2q + 64 \quad \text{and} \quad D(q) = 100 - \frac{q^2}{3}.$$

Determine the consumer and producer surpluses.

*Solution.* The market price and quantity occur when supply and demand are in equilibrium, or $S(q) = D(q)$. Setting these to be equal we get

$$\frac{q^2}{3} + 2q + 64 = 100 - \frac{q^2}{3} \quad \Rightarrow \quad \frac{2q^2}{3} + 2q - 36 = 0. \quad \Rightarrow \quad \frac{2}{3}(q-6)(q+9) = 0.$$

Discarding the negative solution, we have equilibrium when $q = 6$. The corresponding price is $D(q) = S(q) = 88$. Computing the consumer and producer surpluses gives

$$\begin{aligned} \text{CS} &= \int_0^6 [D(q) - 88]\, \mathrm{d}q = \int_0^6 \left[12 - \frac{q^2}{3}\right] \mathrm{d}q = 48. \\ \text{PS} &= \int_0^6 [88 - S(q)]\, \mathrm{d}q = \int_0^q \left[24 - 2q - \frac{q^2}{3}\right] \mathrm{d}q = 84. \end{aligned}$$

■

## 9.4 Exercises

9-1. Find the total area of the regions enclosed by the given curves:

(a) $x = y^2$ and $x = y + 2$
(b) $x = y^2$ and $x + 2y^2 = 3$
(c) $y = x^4 - x^2$ and $y = 1 - x^2$
(d) $y = x\sqrt{1-x^2}$ and $y = x - x^3$
(e) $y = e^x$, $y = e^{3x}$, and $x = 1$
(f) $y = 2^x, y = 5^x$, and $x = 1$

9-2. Find the total area between the given curves:

(a) $y^2 - 4x = 3$ and $4x - y = 3$
(b) $y^2 - x = 2, x - e^y = 0, y = \pm 1$

9-3. Find the total area bounded between $y = x^2 - 2$ and $y = x$.

9-4. Find the number $c$ such that $y = c$ divides the region bounded by $y = x^2$ and $y = 4$ into two subregions of equal area.

9-5. Consider the curves $y = mx$ and $y = x^3$, where $m$ is some positive real number. Find the value of $m$ such that the area bounded by these two curves is exactly 8.

9-6. (a) Consider the box whose vertices are $(-1,1), (-1,0), (1,0)$, and $(1,1)$. Let $L_m$ be the line with slope $m$ that passes through the point $(2,0)$. Find the value of $m \in \mathbb{R}$ which divides the box into two halves, each of equal area.

(b) Repeat the problem above, but replace the box with the triangle whose vertices are $(0,0), (1,0)$, and $(0,1)$;

9-7. Find a positive, continuous function $f$ such that the area between the $x$-axis and the graph of $f$ between 0 and $x$ is exactly $A(x) = x^4$.

9-8. Find the area of the triangles with vertices:

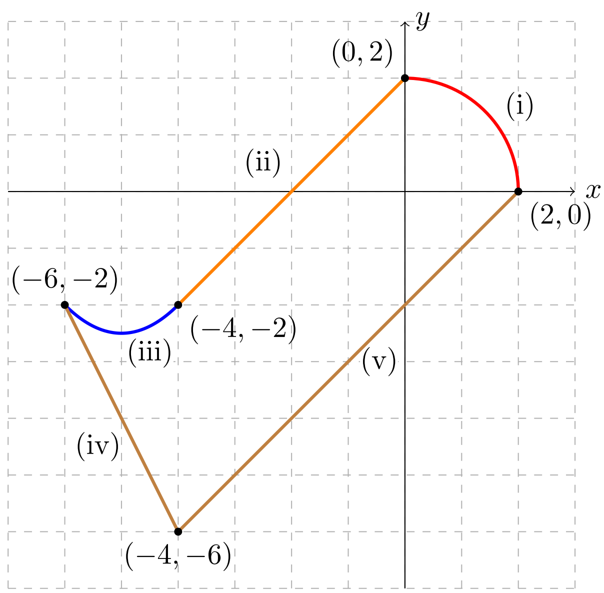
(a) $(0,0), (1,8), (4,3)$
(b) $(-2,5), (0,-3), (5,2)$

9-9. Find the area of the set satisfying $y^2 - x \leq 2$, $x \leq y$, and $x \leq 1$.

9-10. The following graph illustrates the region bounded by:

(i) $x = \sqrt{4-y^2}$
(ii) $y = x+2$
(iii) $y = \frac{1}{2}(x+6)(x+4)-2$
(iv) $y = -2x-14$
(v) $y = x-2$

Calculate the area of the region. Explain your choice of bounds carefully.




9-11. Let $R$ be the region bounded by $y = x^2 - 1$ and $y = 9 - 3x$. Let $S = \{(x, y) : x > 0, y > 0\}$; that is, the portion of the $xy$-plane in the top-right corner. Determine the area of $R \cap S$.

9-12. Let $R$ be the region bounded by the curves $y = 5x^2 - 6$ and $y = 3x^2 + 2$.

(a) Write the area of $R$ as an integral by integrating in the $x$-direction. When setting up the integrals, be sure to explain your choice of bounds.

(b) Write the area of $R$ as an integral by integrating in the $y$-direction. When setting up the integrals, be sure to explain your choice of bounds.

(c) Determine the area of $R$.

9-13. Solve the following improper integrals:

(a) $\displaystyle\int_0^\infty 3^{-x}\,\mathrm{d}x$

(b) $\displaystyle\int_0^\infty xe^{-x^2}\,\mathrm{d}x$

(c) $\displaystyle\int_{-\infty}^0 \frac{1}{5-4x}\,\mathrm{d}x$

(d) $\displaystyle\int_{10}^\infty \frac{1}{x\ln x}\,\mathrm{d}x$

(e) $\displaystyle\int_{-\infty}^\infty e^{-|x|}\,\mathrm{d}x$

(f) $\displaystyle\int_0^\infty xe^{-5x}\,\mathrm{d}x$

9-14. Is there any $p \in \mathbb{R}$ such that $\displaystyle\int_0^\infty \frac{1}{x^p}\,\mathrm{d}x$ converges?

9-15. Determine whether the following integrals converges:

(a) $\displaystyle\int_0^\infty \frac{x+1}{e^x - x}\, \mathrm{d}x$

(b) $\displaystyle\int_1^\infty \frac{1}{x^2+x}\, \mathrm{d}x$

(c) $\displaystyle\int_2^\infty \frac{\sqrt{x+2}}{x+1}\, \mathrm{d}x$

(d) $\displaystyle\int_1^\infty \frac{x+5}{x^3+3x^2-1}\, \mathrm{d}x$

9-16. (a) Suppose \$12,000 is invested and compounded continuously over a span of 5 years at an APR of 4.2%. Determine the value of the annuity after those five years.

(b) If \$20,000 is invested and compounded continuously for four years, at what rate would you need to compound to guarantee a final value of \$100,000?

9-17. Consider the market for hockey pucks, sold by the bucket. We'll let $p$ describe the price per bucket, and $q$ the number of buckets. Let $p = S(q)$ be the *supply curve*, which describes the price $p$ at which a producer is willing to sell $q$ pucks. Similarly, let $p = D(q)$ be the *demand curve*, which describes the price $p$ at which a consumer is willing to purchase $q$ pucks.

(a) *Market equilibrium* occurs when the consumer and producer agree on both the price and quantity to buy and sell, respectively. Suppose

$$S(q) = 4 + \frac{q^2}{36} \quad \text{and} \quad D(q) = 5 - \frac{q^2}{12}.$$

Determine the market equilibrium for this example. Call this point $(p_0, q_0)$.

(b) Suppose you purchase $q_1$ pucks at a price of $p_1 = D(q_1)$. The *consumer surplus* CS is the benefit you received for purchasing at this price, versus a worse price point. Geometrically, this corresponds to the area beneath $p = D(q)$ and above the straight line $p = p_1$ on $[0, q_1]$. Write this as an integral.

(c) Analogously, if a producer can sell $q_2$ pucks at a price $p_2 = S(q_2)$, the *producer surplus* is the benefit received for selling at this price versus a lower price. Geometrically, this corresponds to the area above $p = S(q)$ and below the straight line $p = p_2$ on $[0, q_2]$. Write this as an integral.

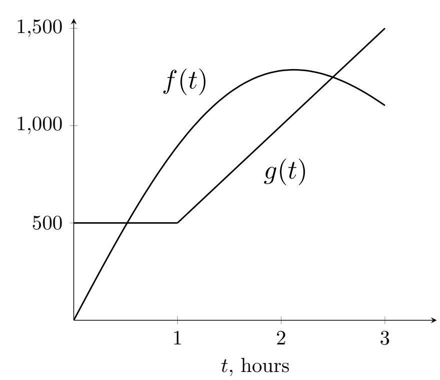
(d) The plots of $S$ and $D$ from part (a) are given below. Draw and shade the consumer and producer surplus. Be sure to clearly indicate which is which.

(e) Assuming $S$ and $D$ are as in part (a), and the market price is $(q_0, p_0)$, compute the consumer and producer surplus.

9-18. Water is flowing into a tank at a rate of $f(t) = 1000te^{-t^2/9}$ L/hr, where $0 \le t \le 3$ is in hours. At the same time, water is exiting the tank at the rate of $g(t)$ L/hr, where

$$g(t) = \begin{cases} 500, & 0 \le t < 1 \\ 500t, & 1 < t \le 3. \end{cases}$$

Suppose that the tank has 750L of water at time $t = 0$.




(You may use the fact that $f$ and $g$ intersect at $t = 0.515$ and $t = 2.498$.)

(a) Compute the total amount of water that *enters* the tank for $0 \leq t \leq 3$.

(b) Compute the total amount of water that *leaves* the tank for $0 \leq t \leq 3$. Then find how much water remains in the tank at $t = 3$ hrs.

(c) Let $A(t)$ denote the amount of water in the tank at time $t$. For $0 \leq t \leq 3$, give the intervals of increasing and decreasing for $A$.

(d) For $0 \leq t \leq 3$, indicate the local maxima and minima for $A$.

# 10 Differential Equations

One of the most important applications of calculus is to differential equations. These are equations which involve a function $y = f(x)$ and its derivatives $y', y''$, et cetera. Such equations often result from understanding a relationship between variables, from which we want to determine the function $f$. For example, the Solow-Swan model satisfies

$$\frac{\mathrm{d}y}{\mathrm{d}t} = sy^\alpha - \delta y, \qquad \text{where } s, \alpha, \delta \text{ are constants.}$$

When you allow things like partial derivatives (which we'll learn shortly), you get partial differential equations, like the infamous Black-Scholes equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

This equation describes the price $V$ of an option as a function of stock price $S$ and time $t$. Here $r$ and $\sigma$ are constants.[4]

Most differential equations do not admit solutions that can be explicitly written down, and so require the use of a computer to solve. We won't worry about such problems now, and instead focus on the few types of differential equations we do know how to solve.

## 10.1 Basic Differential Equations

The simplest and least interesting type of differential equation is one that can be explicitly integrated. For example, consider the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = x^2 + x.$$

To solve for $y$ as a function of $x$, we can integrate both sides and apply the Fundamental Theorem of Calculus

$$\int \frac{\mathrm{d}y}{\mathrm{d}x} = y = \int \left[x^2 + x\right]\, \mathrm{d}x = \frac{1}{3}x^3 + \frac{1}{2}x^2 + C.$$

If we are given an initial condition, such as $y(0) = 5$, we can solve for $C$ explicitly:

$$5 = y(0) = \frac{1}{3}0^3 + \frac{1}{2}0^2 + C = C$$

so $C = 5$ and our solution is $y = x^3/3 + x^2/2 + 5$. Providing a different initial condition changes the solution.

[4]The Black-Scholes equation is actually a *Stochastic Differential Equation*, but stochastic calculus is beyond the scope of this course.

## 10.2 Separable Differential Equations

Basic differential equations are not very interesting, since they're solved by simply integrating. A much more interesting problem is something like

$$\frac{dy}{dx} = y. \tag{10.1}$$

This is a naturally occurring differential equation. It says that the growth of $y$ is proportional to $y$ itself, so when $y$ is small $\frac{dy}{dx}$ is small, and when $y$ is large $\frac{dy}{dx}$ is large. What do we know of that exhibits proportional growth? Remember, here we're thinking of $y$ *as a function of* $x$. We cannot just integrate, since we don't know how to make sense of the term $\int y(x)\,dx$. In this case we can guess what the solution might be. What function do we know of that differentiates to itself?

The first class of solvable differential equations we examine are called *separable*. A differential equation is separable if we can segregate the dependent variable $y$ from the independent variable $x$, treating $\frac{dy}{dx}$ as a fraction. Equation (10.1) is separable, since we can "multiply" both sides by $dx$ to get

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

Applying an integral to each side, we get

$$\begin{aligned} \int \frac{1}{y}\,dy &= \ln(y) && \text{left hand side} \\ &= \int dx = x + C && \text{right hand side,} \end{aligned}$$

so $\ln(y) = x + C$. We want to write $y$ as a function of $x$ though, so exponentiating both sides gives $y = e^{x+C} = Ke^x$. If you guessed that an exponential function was the solution to this ODE, you were right! We can double check our solution by plugging it back into the ODE:

$$\frac{dy}{dx} = Ke^x = y.$$

Just like before, given an initial condition we can determine the value of the constant $K$ explicitly. For example, if $y(0) = 1$ then

$$10 = y(0) = Ke^0 = K$$

so $y(x) = 10e^x$.

> **Separable Differential Equation** Given a differential equation of the form
>
> $$\frac{dy}{dx} = f(x)g(y)$$
>
> segregate all $x$'s and $y$'s on separate sides
>
> $$\frac{1}{g(y)}\,dy = f(x)\,dx$$
>
> and integrate to find $y$ as a function of $x$.

**Example 10.1**

Find the general solution to the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 3y^4x^2.$$

*Solution.* This is a separable equation, which we can write as

$$\frac{1}{y^4}\,\mathrm{d}y = 3x^2\,\mathrm{d}x.$$

Integrating both sides we get

$$\begin{aligned} \int \frac{1}{y^4}\,\mathrm{d}y &= -\frac{1}{3}\frac{1}{y^3} && \text{left hand side} \\ &= \int 3x^2\,\mathrm{d}x = x^3 + C && \text{right hand side.} \end{aligned}$$

We solve for $y$ by taking the reciprocal and the cube root to get

$$y^3 = -\frac{1}{3x^3+C} \qquad \Rightarrow \qquad y = -\frac{1}{\sqrt[3]{3x^3+C}}.$$

■

**Example 10.2**

Determine the unique function $y(x)$ which satisfies the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = e^y(3x-5), \qquad \text{with} \qquad y(0) = 1.$$

*Solution.* This is a separable differential equation, which we can split as

$$e^{-y}\,\mathrm{d}y = (3x-5)\,\mathrm{d}x.$$

Integrating each side gives

$$\begin{aligned} \int e^{-y}\,\mathrm{d}y &= -e^{-y} && \text{left hand side} \\ &= \int (3x-5)\,\mathrm{d}x && \text{right hand side} \\ &= \frac{3}{2}x^2 - 5x + C. \end{aligned}$$

We solve for $y$ by taking the logarithm of both sides:

$$e^{-y} = -\frac{3}{2}x^2 - 5x + C \qquad \Rightarrow \qquad y = -\ln\left(-\frac{3}{2}x^2 + 5x + C\right).$$

To find the value of $C$, we use the initial condition $y(0) = 1$

$$-1 = y(0) = \ln(C)$$

so $C = e^{-1}$, and the solution is $y(x) = \ln(-3x^2/2 - 5x + e^{-1})$. ■

## 10.3 Linear Differential Equations

Linear differential equations are those of the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} + f(x)y = g(x)$$

for some constant $A$ and function $f$. Such equations are not separable, since multiplying by $\mathrm{d}x$ gives

$$\mathrm{d}y + \underbrace{f(x)y\,\mathrm{d}x}_{(\dagger)} = g(x)\,\mathrm{d}x,$$

and the ($\dagger$) term involves both $x$'s and $y$'s. Instead, if $F$ is an anti-derivative of $f$, think about differentiating the product $ye^{F(x)}$, which gives

$$\frac{\mathrm{d}y}{\mathrm{d}x}e^{F(x)} + f(x)ye^{F(x)} = e^{F(x)}\left[\frac{\mathrm{d}y}{\mathrm{d}x} + f(x)y\right].$$

The term in square brackets is precisely the left hand side of our linear differential equation. This suggests that if we multiply our entire differential equation by $e^{F(x)}$ – called the *integrating factor* – we can simplify the left hand side into something which we can integrate.

**Linear Differential Equations:** Given a differential equation of the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} + f(x)y = g(x)$$

multiply everything by $I(x) = e^{\int f(x)\,\mathrm{d}x}$, so that

$$\frac{\mathrm{d}y}{\mathrm{d}x}I(x) + f(x)yI(x) = g(x)I(x) \quad \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}x}\left[yI(x)\right] = g(x)I(x).$$

$$y = \frac{1}{I(x)}\int g(x)I(x)\,\mathrm{d}x$$

**Example 10.3**

Determine a solution to the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} + 3y = 9, \qquad \text{where} \qquad y(0) = 6.$$

*Solution.* Our integrating factor is $I = e^{3x}$, which we multiply into the differential equation to get

$$\frac{\mathrm{d}y}{\mathrm{d}x}e^{3x} + 3e^{3x}y = 9e^{3x} \quad \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}x}\left[ye^{3x}\right] = 9e^{3x}.$$

Integrating both sides

$$ye^{3x} = \int 9e^{3x}\,\mathrm{d}x = 3e^{3x} + C \quad \Rightarrow \quad y = 3 + Ce^{-3x}.$$

Subbing in $y(0) = 6$ gives $6 = y(0) = 3 + C$ so $C = 3$, and our final solution is $y(x) = 3 + 3e^{-3x}$. Again, we can check this by substituting it back into our differential equation. Since $y'(x) = -9e^{-3x}$ we get

$$\frac{\mathrm{d}y}{\mathrm{d}x} + 3y = -9e^{-3x} + 3\left[3 + 3e^{-3x}\right] = -9e^{-3x} + 9 + 9e^{-3x} = 9. \qquad \blacksquare$$

**Example 10.4**

Solve the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} + 2xy = 5x \qquad \text{where} \qquad y(0) = 1/2.$$

*Solution.* Our integrating factor is $I(x) = e^{\int 2x\,\mathrm{d}x} = e^{x^2}$. Multiplying through gives

$$e^{x^2}\frac{\mathrm{d}y}{\mathrm{d}x} + 2xe^{x^2}y = 5xe^{x^2} \quad \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}x}\left[ye^{x^2}\right] = 5xe^{x^2}.$$

Integrating both sides – using substituting for the right hand side – we get

$$ye^{x^2} = \frac{5}{2}e^{x^2} + C,$$

which we solve for $y$ to get $y(x) = 5/2 + Ce^{-x^2}$. Subbing in our initial condition,

$$\frac{1}{2} = y(0) = \frac{5}{2} + C \quad \Rightarrow \quad C = -2,$$

so our final solution is $y(x) = 5/2 - 2e^{-x^2}$. $\blacksquare$

## 10.4 Second Order Differential Equations

The separable and linear ordinary differential equations seen above are all *first order*, in that they only involve the first derivative $y'$ and $y$. A *second order* differential equation involves $y'', y'$, and $y$. Solving such systems involves a lot of work, but we're going to restrict ourselves to equations of the form

$$a\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + b\frac{\mathrm{d}y}{\mathrm{d}x} + cy = 0. \tag{10.2}$$

This looks a lot like a quadratic equation $ax^2+bx+c = 0$. Moreover, since we need a function $y(x)$ whose first and second derivatives all cancel, we're inspired to try a function of the form $y = e^{mx}$. In this case we have

$$y = e^{mx}, \quad \frac{\mathrm{d}y}{\mathrm{d}x} = me^{mx}, \quad \text{and} \quad \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = m^2e^{mx}.$$

Substituting this into (10.2) gives

$$0 = am^2e^{mx} + bme^{mx} + ce^{mx} = e^{mx}\left[am^2 + bm + c = 0\right].$$

Since $e^{mx} \neq 0$ for all $x$, it must be the case that $am^2 + bm + c = 0$, meaning we can use the quadratic formula to determine the values of $m$.

**Second Order Differential Equation:** Given a differential equation of the form

$$a\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + b\frac{\mathrm{d}y}{\mathrm{d}x} + cy = 0$$

the solution is

1. $y(x) = Ae^{r_1 x} + Be^{r_2 x}$ if $r_1$ and $r_2$ are distinct solutions to $am^2 + bm + c = 0$.
2. $y(x) = (A + Bx)e^{rx}$ if $r$ is the only solution to $am^2 + bm + c = 0$.

Note that second order differential equations require *two* initial conditions to uniquely identify the solution.

**Example 10.5**

Find the solution to the differential equation

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + \frac{\mathrm{d}y}{\mathrm{d}x} - 2y = 0 \quad \text{where} \quad y(0) = 6 \quad \text{and} \quad y'(0) = 0.$$

*Solution.* The coefficients are $a = 1, b = 1$, and $c = -2$, giving the quadratic equation $0 = m^2 + m - 2 = (m+2)(m-1)$ showing that $r_1 = -2, r_2 = 1$ are the roots. Thus our solution is of the form $y(x) = Ae^{-2x} + Be^{x}$. Using our initial conditions

$$\begin{aligned} 6 &= y(0) = A + B \\ 0 &= y'(0) = \left[-2Ae^{-2x} + Be^{x}\right]_{x=0} = -2A + B \end{aligned}$$

Solving this system gives $A = 2$ and $B = 4$, so the final solution is $y(x) = 2e^{-2x} + 4e^{x}$. ■

**Example 10.6**

Find the solution to the ordinary differential equation

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} - 4\frac{\mathrm{d}y}{\mathrm{d}x} + 4y = 0 \quad \text{where} \quad y(0) = 1 \quad \text{and} \quad y'(0) = -1.$$

*Solution.* Here $a = 1, b = -4, c = 4$ so we find the solutions to $0 = m^2 - 4m + 4 = (m-2)^2$. We get a single root $r = 2$, so our general solution is of the form $y(x) = (A + Bx)e^{2x}$. The initial conditions give

$$\begin{aligned} 1 &= y(0) = A \\ -1 &= y'(0) = \left[(2A + B)e^{2x} + 2Bxe^{2x}\right]_{x=0} = 2A + B, \end{aligned}$$

so $A = 1$ and $B = -3$. The final solution is thus $y(x) = (1 - 3x)e^{2x}$. ■

## 10.5 Exercises

10-1. Find the general solution to each of the following differential equations:

(a) $\dfrac{dy}{dx} = 4x^3 e^{2y}$

(b) $\dfrac{dy}{dx} = -2x^5 y^7$

(c) $\dfrac{dy}{dx} = \dfrac{2y+3}{x-1}$

(d) $\dfrac{dy}{dx} = e^{3x-y}$

(e) $(1+x^2)\dfrac{dy}{dx} + xy = 0$

(f) $x\dfrac{dx}{dy} + x^2 y = 0$

10-2. Find the unique solution to each given differential equation:

(a) $\dfrac{dy}{dx} = \dfrac{y^3}{x^2}$ where $y(1) = -1/2$.

(b) $e^{y^2}\dfrac{dy}{dx} = \dfrac{x}{y}$ where $y(0) = 0$

(c) $y^2\dfrac{dy}{dx} - xe^x = 0$ where $y(0) = \sqrt[3]{3}$.

10-3. Find general solutions to the following differential equations:

(a) $\dfrac{dy}{dx} + 4y = x$

(b) $\dfrac{dy}{dx} + 4xy = x$

(c) $x\dfrac{dy}{dx} + y = 3x + e^{2x}$

(d) $\dfrac{dy}{dx} + \dfrac{y}{x} = \dfrac{1}{2+x^2}$

10-4. Repeat Question (3) with the following initial conditions:

(a) $y(0) = 1$

(b) $y(0) = -3$

(c) $y(1) = e$

(d) $y(1) = \ln(3)$

10-5. Solve each given differential equation:

(a) $y'' - 5y' + 6y = 0$

(b) $y'' - y = 0$ with $y(0) = 2$ and $y'(0) = 1$

(c) $y'' + 2y' - 24y = 0$

(d) $y'' + 4y' + 4y = 0$ with $y(0) = 4$ and $y'(0) = -1$

# 11 Multivariable Calculus

In this section, we superficially scratch the surface of functions of several variables, and how calculus extends to these functions. This situation is *significantly* more complicated. In general, functions now accept multiple parameters, such as $f(x, y, z) = xy + x^2 - z^2$. To evaluate such a function, we need to be given a triple, such as $(x, y, z) = (1, 3, 5)$ which can then be substituted into $f$ to get

$$f(1, 3, 5) = (xy + x^2 - z^2)_{x=1,y=3,z=5} = (1)(3) + (1)^2 - (5)^2 = 3 + 1 - 25 = -21.$$

In the special case of a two-parameter function, we can still visualize it in terms of a three-dimensional graph $z = f(x, y)$, as pictured in Figure 11.1.

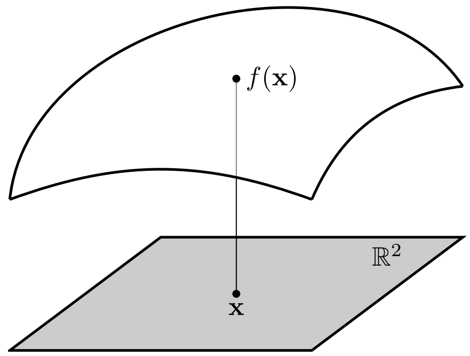



Figure 11.1: A function $f : \mathbb{R}^n \to \mathbb{R}$ can be visualized in terms of its graph.

When referring to the input arguments of a function, there are two conventions to which we'll adhere. The first is the usual alphabetical naming of variables, such as $f(x, y, z)$. Here it is clear that $x$ is the first argument, $y$ is the second, and $z$ is the third. However, we could also write $f(x_1, x_2, x_3)$, where $x_i$ is the $i$th variable.

## 11.1 Partial Derivatives

It's no longer possible to describe the "slope" of a plane, so the best we can do is talk about the slope of a line in a particular direction. This brings about the definition of a partial derivative.

**Definition 11.1**

Write $(x_1, \ldots, x_n)$ to denote the coordinates of $\mathbb{R}^n$. If $f$ is a function of $n$-variables, we define the partial derivative of $f$ with respect to $x_i$ at $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{R}^n$ as

$$\frac{\partial f}{\partial x_i}(\mathbf{a}) = \lim_{h \to 0} \frac{f(a_1, \ldots, a_i + h, \ldots, a_n) - f(a_1, \ldots, a_n)}{h}.$$

That is, $\frac{\partial f}{\partial x_i}$ is the one-variable derivative of $f(x_1, \ldots, x_n)$ with respect to $x_i$, where all other variables are held constant.

The partial derivative of $f$ with respect to $x$ is the instantaneous rate of change of $f$ if we moved in the $x$-direction only.

**Example 11.2**

If $f(x, y) = x^2 + xy$, find $\dfrac{\partial f}{\partial x}(\mathbf{a})$ where $\mathbf{a} = (1, -1)$.

*Solution.* Substituting into the definition of a partial derivative, we have

$$\begin{aligned}\frac{\partial f}{\partial x}(1,-1) &= \lim_{h\to 0}\frac{f(1+h,-1)-f(1,-1)}{h} = \lim_{h\to 0}\frac{[(1+h)^2+(1+h)(-1)]-[(1)^2+(1)(-1)]}{h}\\ &= \lim_{h\to 0}\frac{[h^2+2h+1-h-1]}{h} = \lim_{h\to 0}\frac{h^2+h}{h} = 1.\end{aligned}$$

■

Just as with single variable derivatives, we can turn partial derivatives into functions by evaluating them in general. For example, using $f(x,y)=x^2+xy$ from Example 11.2, and evaluating at a generic point $(x,y)$, we get

$$\begin{aligned}\frac{\partial f}{\partial x}(x,y) &= \lim_{h\to 0}\frac{f(x+h,y)-f(x,y)}{h} = \lim_{h\to 0}\frac{[(x+h)^2+(x+h)y]-[x^2+xy]}{h}\\ &= \lim_{h\to 0}\frac{x^2+2xh+h^2+xy+yh-x^2-xy}{h} = \lim_{h\to 0}\frac{2xh+h^2+yh}{h}\\ &= 2x+y.\end{aligned}$$

If we plug in $(x,y)=(1,-1)$ we get $\frac{\partial f}{\partial x}(1,-1)=2(1)+(-1)=1$, which agrees with what we found earlier.

Without having to reinvent derivative rules, look again at the definition of the partial derivative. It looks exactly like a normal derivative in one of the variables, with all the other variables held constant. Hence we can compute partial derivatives by adhering to the same philosophy: Treat everything other than the variable you're differentiating as a constant.

**Example 11.3**

Determine the partial derivatives of the function $f(x,y,z)=xy+\ln(x^2z)+z^{-2}e^y$.

*Solution.* When computing the partial derivative with respect to $x_i$, we treat all other variables as constants. Hence

$$\begin{aligned}\frac{\partial f}{\partial x} &= y+\frac{2xz}{x^2z} = y+\frac{2}{x}\\ \frac{\partial f}{\partial y} &= x+\frac{e^y}{z^2}\\ \frac{\partial f}{\partial z} &= \frac{1}{z}-\frac{2e^y}{z^3}.\end{aligned}$$

■

If $f$ is a function of $n$-variables and all the partial derivatives exist, the *gradient* of $f$ is defined to be

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}),\frac{\partial f}{\partial x_2}(\mathbf{x}),\ldots,\frac{\partial f}{\partial x_n}(\mathbf{x})\right).$$

It can be quite cumbersome to write $\frac{\partial f}{\partial x_i}$, so we will often interchange it with any of the following when it is unambiguous:

$$\frac{\partial f}{\partial x_i},\quad \partial_{x_i}f,\quad \partial_i f,\quad f_{x_i},\quad f_i.$$

This will be particularly convenient when we start taking higher order partial derivatives.

**Example 11.4**

Consider the function $f(x, y) = xye^{x^2 y}$. Find the partial derivatives of $f$.

*Solution.* Let's start with the $x$-partial derivative. Treating $y$ as a constant we get

$$f_x(x, y) = ye^{x^2 y} + xy(2xye^{x^2 y}) = ye^{x^2 y} + 2x^2 y^2 e^{x^2 y} = ye^{x^2 y}(1 + 2x^2 y).$$

The $y$-partial is similar, with

$$f_y(x, y) = xe^{x^2 y} + x^3 ye^{x^2 y} = xe^{x^2 y}(1 + x^2 y).$$ ∎

**Example 11.5**

Consider the function $f(x, y) = 10 + x^3 + y^3 - 3xy$. Find all $\mathbf{x} \in \mathbb{R}^2$ such that $\nabla f(\mathbf{x}) = 0$.

*Solution.* The gradient consists of the partial derivatives, which we compute and set to $\mathbf{0}$.

$$\nabla f(x, y) = (3x^2 - 3y, 3y^2 - 3x) = (0, 0).$$

This gives a system of equations $x^2 = y$ and $y^2 = x$. If we plug the first equation into the second,

$$x = y^2 = (x^2)^2 = x^4 \quad \Rightarrow \quad x(x^3 - 1) = 0,$$

which has solution $x = 0$ and $x = 1$. When $x = 0$ we get $y = 0$ and when $x = 1$ we get $y = 1$, so the points where the gradient is zero are $(0, 0)$ and $(1, 1)$. ∎

## 11.2 Applications of Partial Derivatives

Just as in the single variable case, slapping a derivative on a quantity is economically known as the marginal value of that quantity. For example, if $C(x, y)$ describes the joint cost for a manufacturer to produce two products with quantities $x$ and $y$, then

$$\frac{\partial C}{\partial x} \quad \text{and} \quad \frac{\partial C}{\partial y}$$

represent the marginal cost to produce $x$ and the marginal cost to produce $y$ respectively. Intuitively, these represent the rate of change of $x$ and $y$ if the other variable is held constant.

**Example 11.6**

Suppose the productivity of your company is measured in terms of capital $k$ and labour $\ell$, and is described by

$$P(k, \ell) = \frac{k\ell}{3k + 5\ell}.$$

Find the marginal productivity functions.

*Solution.* Taking partial derivatives with respect to $k$ and $\ell$ gives

$$\begin{aligned}\frac{\partial P}{\partial k} &= \frac{\ell(3k+5\ell)-3(k\ell)}{(3k+5\ell)^2} = \frac{5\ell^2}{(3k+5\ell)^2}\\ \frac{\partial P}{\partial \ell} &= \frac{k(3k+5\ell)-5(k\ell)}{(3k+5\ell)^2} = \frac{3k^2}{(3k+5\ell)^2}.\end{aligned}$$

■

More interesting is that we can now talk about relationships between objects. For example, suppose we examine the demand functions for Android and iPhone products as a function of their corresponding prices: $q_A = D_A(p_A, p_I), q_I = D_I(p_A, p_I)$. Because these two items compete with one another, one would expect that

$$\frac{\partial q_A}{\partial p_I} > 0 \quad \text{and} \quad \frac{\partial q_I}{\partial p_A} > 0.$$

Think about what this means: $\frac{\partial q_A}{\partial p_I}$ describes how the demand for Android phones changes as a function of iPhone pricing. The fact that the partial derivative is positive means that this is increasing in $p_I$; that is, as iPhones become more expensive, the demand for Android phones should increase. A similar explanation works for $\frac{\partial q_I}{\partial p_A}$ as well. These are *competitive products.*

On the other hand, consider two manufacturers, one who makes shoes and one that makes shoe laces. Their demand functions similarly depend on the prices of one another, with $q_S = D(p_S, p_L)$ and $q_L = D(p_S, p_L)$. Unlike the Android and iPhone example above, we expect these demand curves to satisfy

$$\frac{\partial q_S}{\partial p_L} < 0 \quad \text{and} \quad \frac{\partial q_L}{\partial p_S} < 0.$$

Once again, $\frac{\partial q_S}{\partial p_L}$ describes the demand for shoes as a function of lace costs. The fact that it is negative means that as shoe lace costs become larger, the demand for shoes decreases. The same explanation holds for $\frac{\partial q_L}{\partial p_s}$. These are said to be *complementary products.*

**Example 11.7**

Two products $A$ and $B$ have demand functions

$$q_A = e^{-p_A+p_B} \quad \text{and} \quad q_B = \frac{16}{p_A^2 p_B^2}.$$

Determine whether these products are complementary, competitive, or neither.

*Solution.* Taking partial derivatives we get

$$\frac{\partial q_A}{\partial p_B} = -e^{-(p_A+p_B)} \quad \text{and} \quad \frac{\partial q_B}{\partial p_A} = \frac{-32}{p_A^3 p_B^2}.$$

Certainly $\frac{\partial q_A}{\partial p_B}$ is always negative, while $\frac{\partial q_B}{\partial p_A}$ is only positive if $p_A < 0$, which is nonsense as prices cannot be negative. Hence both derivatives are negative, showing that $A$ and $B$ are complementary products. ■

## 11.3 Higher-Order Partial Derivatives

For differentiable functions of one variable, a lot of information about $f$ could be inferred not only from its first derivative $f'$, but from its higher order derivatives $f^{(n)}$. For example, if $f$ represents some physical quantity such as position as a function of time, we know that $f'$ is its velocity, $f''$ is its acceleration, and $f^{(3)}$ is its jerk. This means that the higher-order derivatives are essential when modelling differential equations. We used higher order derivatives when approximating functions with polynomials, and we exploited the second derivative test to determine optimality of critical points. All of these applications and more will extend to functions of $n$-variables as well.

The first step is second-order derivatives; that is, to differentiate a function twice. Interestingly though, we now have many different ways of computing a second derivative. For example, if $f$ is a function of two variables, then there are four possible second derivatives:

$$\partial_{xx} f = \frac{\partial}{\partial x}\left[\frac{\partial f}{\partial x}\right], \qquad \partial_{xy} f = \frac{\partial}{\partial x}\left[\frac{\partial f}{\partial y}\right], \qquad \partial_{yx} f = \frac{\partial}{\partial y}\left[\frac{\partial f}{\partial x}\right], \qquad \partial_{yy} f = \frac{\partial}{\partial y}\left[\frac{\partial f}{\partial y}\right].$$

The terms $\partial_{xx} f, \partial_{yy} f$ are called *pure partial derivatives*, while $\partial_{xy} f, \partial_{yx} f$ are called *mixed partial derivatives*. In general, given a function of $n$-variables, there are $n^2$ different second-order partial derivatives.

**Example 11.8**

Determine the second-order partial derivatives of the function $f(x, y) = e^{xy} + x^2 \ln(y)$.

*Solution.* This is a matter of straightforward computation. The first order partial derivatives are given by

$$\frac{\partial f}{\partial x} = y e^{xy} + 2x \ln(y) \quad \text{and} \quad \frac{\partial f}{\partial y} = x e^{xy} + \frac{x^2}{y}.$$

To compute the second order partials, we treat each of the first order partials as functions of $x$ and $y$ and repeat the process:

$$\begin{aligned} \partial_{xx} f(x,y) &= y^2 e^{xy} + 2\ln(y) & \qquad \partial_{xy} f(x,y) &= e^{xy} + xy e^{xy} + \tfrac{2x}{y} \\ \partial_{yx} f(x,y) &= e^{xy} + xy e^{xy} + \tfrac{2x}{y} & \qquad \partial_{yy} f(x,y) &= x^2 e^{xy} - \tfrac{x^2}{y^2}. \end{aligned}$$

Interestingly, note that $\partial_{yx} f = \partial_{xy} f$. ■

**Example 11.9**

Determine the second-order partial derivatives of the function $f(x, y, z) = e^{xy+xz}$.

*Solution.* The first order partial derivatives are given by

$$\partial_x f(\mathbf{x}) = (y+z) e^{xy+xz}, \quad \partial_y f(\mathbf{x}) = x e^{xy+xz}, \quad \text{and} \quad \partial_z f(\mathbf{x}) = x e^{xy+xz}.$$

There are 9 second order partials:

$$\begin{array}{lll} \partial_{xx} f(\mathbf{x}) = (y+z)^2 e^{xy+xz} & \partial_{xy} f(\mathbf{x}) = (1+xy+xz)e^{xy+xz} & \partial_{xz} f(\mathbf{x}) = (1+xy+xz)e^{xy+yz} \\ \partial_{yx} f(\mathbf{x}) = (1+xy+xz)e^{xy+yz} & \partial_{yy} f(\mathbf{x}) = x^2 e^{xy+xz} & \partial_{yz} f(\mathbf{x}) = x^2 e^{xy+xz} \\ \partial_{zx} f(\mathbf{x}) = (1+xy+xz)e^{xy+xz} & \partial_{zy} f(\mathbf{x}) = x^2 e^{xy+xz} & \partial_{zz} f(\mathbf{x}) = x^2 e^{xy+xz} \end{array}.$$

The table exhibits a great deal of symmetry. ■

The fact that $\partial_{zz} f = \partial_{yy} f$ in Example 11.9 is a consequence of the symmetry of the function $f(x, y, z) = e^{xy+xz}$: Interchanging $y$ and $z$ never changes the value of the function. However, somewhat more surprising is that in both of the previous two examples our mixed partial derivatives were the same. It turns out that this is a fairly common occurrence.

**Theorem 11.10: Clairut's Theorem**

Let $f$ be an $n$-variable function and $\mathbf{a} \in \mathbb{R}^n$ a point. Let $i, j \in \{1, \ldots, n\}$ with $i \neq j$. If $\partial_{ij} f(\mathbf{a})$ and $\partial_{ji} f(\mathbf{a})$ both exist and are continuous near $\mathbf{a}$, then $\partial_{ij} f(\mathbf{a}) = \partial_{ji} f(\mathbf{a})$.

We often collect the second order derivatives into a matrix $H_f(\mathbf{x}) = \left[\frac{\partial^2 f}{\partial x_i x_j}\right]_{ij}$ called the *Hessian matrix*. From Clairut's theorem, the Hessian matrix is symmetric, in the sense that $H_f(\mathbf{x})^T = H_f(\mathbf{x})$.

**Example 11.11**

Evaluate the Hessian matrix from the function $f(x, y) = e^{xy} + x^2 \ln(y)$ at the point $(x, y) = (3, 1)$.

*Solution.* We computed the second derivatives in Example 11.8, giving

$$H_f(x, y) = \begin{bmatrix} y^2 e^{xy} + 2\ln(y) & e^{xy} + xye^{xy} + 2x/y \\ e^{xy} + xye^{xy} + 2x/y & x^2 e^{xy} - x^2/y^2 \end{bmatrix}$$

Evaluating at the point $(x, y) = (3, 1)$ we get

$$H_f(3, 1) = \begin{bmatrix} e^3 & 4e^3 + 6 \\ 4e^3 + 6 & 9e^3 - 9 \end{bmatrix}.$$ ■

### 11.3.1 The Chain Rule

Despite having constantly and consistently cautioned against treating differentials as fractions, there have not been too many instances to date where ignoring this advice could have caused any damage. Here at last our efforts will be vindicated, as I'll show you some of the deeper subtleties in using higher-order partial derivatives in conjunction with the chain rule.

Let's start with a simple but general example. To make a point, we will write all partial derivatives using Leibniz notation. Let $u = f(x, y)$ and suppose that both $x, y$ are functions of

$(s,t)$; that is, $x = g(s,t)$ and $y = h(s,t)$. The derivative $\frac{\partial u}{\partial s}$ is determined by finding all the ways $s$ is related to $u$. This can be summarized in a nice tree:

$$
\begin{tikzcd}
 & & u & & \\
 & x \arrow[ur] & & y \arrow[ul] & \\
s \arrow[ur] & t \arrow[u] & & s \arrow[u] & t \arrow[ul]
\end{tikzcd}
$$

Hence
$$\frac{\partial u}{\partial s} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial s}.$$
Similarly,
$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial t}.$$
Just like the one-dimensional chain rule, you have to be careful about where these are evaluated. Strictly speaking, the correct expression is
$$\left.\frac{\partial u}{\partial t}\right|_{(s,t)} = \left.\frac{\partial u}{\partial x}\right|_{(x(s,t),y(s,t))} \left.\frac{\partial x}{\partial t}\right|_{(s,t)} + \left.\frac{\partial u}{\partial y}\right|_{(x(s,t),y(s,t))} \left.\frac{\partial y}{\partial t}\right|_{(s,t)}.$$

**Example 11.12**

Suppose $z = e^{x+y}$, $x = 2t^2 + 4$, and $y = 1 - t^3$. Explicitly write $z$ as a function of $t$ and compute the derivative of $z$ with respect to $t$. Verify your result using the chain rule.

*Solution.* The expression $x + y$ evaluates to $x + y = -t^3 + 2t^2 + 5$, so that
$$z = e^{x+y} = e^{-t^3+2t^2+5} \qquad \Rightarrow \qquad \frac{\mathrm{d}z}{\mathrm{d}t} = (-3t^2 + 4t)e^{-t^3+2t+5}.$$
On the other hand, the chain rule tells us that
$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\partial z}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial z}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} = e^{x+y}(4t) + e^{x+y}(-3t^2) = (-3t^2 + 4t)e^{x+y}$$
which is the same expression once $x + y$ is evaluated. ■

In the above example, it didn't seem like we saved much time employing the chain rule. Let's increase the difficulty a bit.

**Example 11.13**

Let $z = xye^x$, $x = s^2 + t^2$ and $y = 2st$. Determine $\dfrac{\partial z}{\partial t}$ at $(1,1)$.

*Solution.* I'm not going to compute the derivative explicitly, but were we to substitute $x$ and $y$ into $z$ we would get

$$z = 2st(s^2+t^2)e^{s^2+t^2}.$$

That's not a pretty expression. If you would like to compute $\frac{\partial z}{\partial t}$ as an exercise, go ahead and check that it confirms what we find below.

Instead, let's use the chain rule. We know that $x(1,1) = 2$ and $y(1,1) = 2$ as well, so that

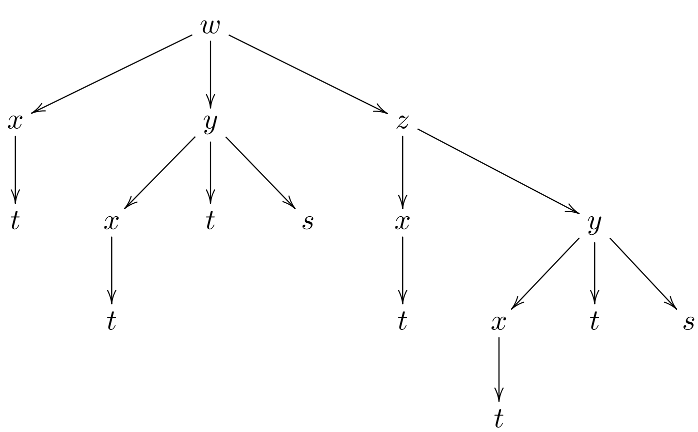
$$\begin{aligned}
\left.\frac{\partial z}{\partial t}\right|_{(1,1)} &= \left.\frac{\partial z}{\partial x}\right|_{(2,2)} \left.\frac{\partial x}{\partial t}\right|_{(1,1)} + \left.\frac{\partial z}{\partial y}\right|_{(2,2)} \left.\frac{\partial y}{\partial t}\right|_{(1,1)} \\
&= \left[ye^x + xye^x\right]_{(2,2)} [2t]_{(1,1)} + [xe^x]_{(2,2)} [2s]_{(1,1)} \\
&= 12e^2 + 4e^2 = 16e^2.
\end{aligned}$$

■

**Example 11.14**

Suppose that $w = f(x, y, z), x = g(t), y = h(x, t, s)$, and $z = r(x, y)$ are all differentiable functions. Use the chain rule to determine $\dfrac{\partial w}{\partial t}$.

*Solution.* We can visualize this with the following tree, where the non-black colours indicate which paths we traverse at each level.




Tracing these paths, we get

$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial y}\frac{\partial y}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial y}\frac{\partial y}{\partial t}.$$

■

The real nastiness comes if you want to take second derivatives. For example, $\partial_{ss}u$ is computed as

$$\frac{\partial^2 u}{\partial s^2} = \frac{\partial}{\partial s}\left[\frac{\partial u}{\partial s}\right] = \frac{\partial}{\partial s}\left[\frac{\partial u}{\partial x}\frac{\partial x}{\partial s}\right] + \frac{\partial}{\partial s}\left[\frac{\partial u}{\partial y}\frac{\partial y}{\partial s}\right].$$

Now realize that since $u = f(x, y)$ is a function of $x$ and $y$, $\frac{\partial u}{\partial x}$ is also a function of $(x, y)$. Thus to differentiate this function with respect to $s$, we must once again use the chain rule. Thus looking at only the first summand, we have

$$\begin{aligned}
\frac{\partial}{\partial s}\left[\frac{\partial u}{\partial x}\frac{\partial x}{\partial s}\right] &= \left[\frac{\partial}{\partial s}\frac{\partial u}{\partial x}\right]\frac{\partial x}{\partial s} + \frac{\partial u}{\partial x}\frac{\partial^2 x}{\partial s^2} && \text{product rule} \\
&= \left[\frac{\partial^2 u}{\partial x^2}\frac{\partial x}{\partial s} + \frac{\partial^2 u}{\partial x \partial y}\frac{\partial y}{\partial s}\right]\frac{\partial x}{\partial s} + \frac{\partial u}{\partial x}\frac{\partial^2 x}{\partial s^2} && \text{chain rule} \\
&= \frac{\partial^2 u}{\partial x^2}\left[\frac{\partial x}{\partial s}\right]^2 + \frac{\partial^2 u}{\partial x \partial y}\frac{\partial y}{\partial s}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial x}\frac{\partial^2 x}{\partial s^2}.
\end{aligned}$$

What a mess! A similar computation on the second summand yields

$$\frac{\partial}{\partial s}\left[\frac{\partial u}{\partial y}\frac{\partial y}{\partial s}\right] = \frac{\partial^2 u}{\partial y^2}\left[\frac{\partial y}{\partial s}\right]^2 + \frac{\partial^2 u}{\partial x \partial y}\frac{\partial y}{\partial s}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial y}\frac{\partial^2 y}{\partial s^2}.$$

Putting everything together:

$$\frac{\partial^2 u}{\partial s^2} = \frac{\partial^2 u}{\partial x^2}\left[\frac{\partial x}{\partial s}\right]^2 + \frac{\partial^2 u}{\partial y^2}\left[\frac{\partial y}{\partial s}\right]^2 + 2\frac{\partial^2 u}{\partial x \partial y}\frac{\partial y}{\partial s}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial x}\frac{\partial^2 x}{\partial s^2} + \frac{\partial u}{\partial y}\frac{\partial^2 y}{\partial s^2}. \tag{11.1}$$

This is only a single partial derivative. The same procedure must also be used to compute $\partial_{xy} u$ and $\partial_{yy} u$.

**Exercise:** Hurt your brain a little bit more! Let $u = f(x, y, s)$ and $x(s, t)$ and $y(s, t)$. Now determine $\partial_{ss} u$.

## 11.4 Optimization

When dealing with differentiable real-valued functions of a single variable we had a standard procedure for determining maxima and minima. This amounted to checking critical points on the interior $(a, b)$ and then checking the boundary points. The necessity of checking the boundary separately arose from the non-differentiability of the function at the boundary. In the multiple dimension regime, we will now be looking at functions of $n$-variables. Once again, we will use differentiability to establish a necessary condition for extrema to occur on the interior, and check the boundary separately. However, unlike the former example where the boundary consisted of two points $\{a, b\}$, in multiple dimensions our boundaries become much larger. This will necessitate an entirely different approach to determining maxima on the boundary.

For now, we recall the definition of what it means to be a local maximum and minimum.

**Definition 11.15**

Let $f$ be a function on $\mathbb{R}^n$:

1. We say that $\mathbf{a} \in \mathbb{R}^n$ is a *local maximum* of $f$ if $f(\mathbf{x}) \leq f(\mathbf{a})$ for all $\mathbf{x}$ near $\mathbf{a}$.
2. We say that $\mathbf{a} \in \mathbb{R}^n$ is a *local minimum* of $f$ if$f(\mathbf{x}) \geq f(\mathbf{a})$ for all $\mathbf{x}$ near $\mathbf{a}$.

When $n = 1$ this is exactly our definition of a maximum/minimum point in $\mathbb{R}$.

### 11.4.1 Critical Points

**Definition 11.16**

If $f$ is a differentiable function in $\mathbb{R}^n$, we say that $\mathbf{c} \in \mathbb{R}^n$ is a *critical point* of $f$ if $\nabla f(\mathbf{c}) = 0$. If $\mathbf{c}$ is a critical point, we say that $f(\mathbf{c})$ is a *critical value*. All points which are not critical are termed *regular points*.

We see that the above definition of a critical point agrees with the our usual definition when $n = 1$; namely, that $f'(c) = 0$.

**Example 11.17**

Determine the critical points of the following functions:

$$f(x, y) = x^3 + y^3, \qquad g(x, y, z) = xy + xz + x$$

*Solution.* The gradient of $f$ is easily determined to be $\nabla f(x, y) = (3x^2, 3y^2)$. Setting this to be $(0, 0)$ implies that $3x^2 = 0 = 3y^2$ so that the only critical point is $(x, y) = (0, 0)$. For the function $g$ we compute $\nabla g(x, y, z) = (y + z + 1, x, x)$. Setting this equal to zero implies that $x = 0$ while $y + z + 1 = 0$. Thus there is an entire line worth of critical points. ■

Notice that critical points do not need to be isolated: one can have entire curves or planes represent critical points. The important property of critical points is that they give a schema for determining when a point is a maximum or minimum, through the following theorem:

**Proposition 11.18**

If $f$ is differentiable on a set $U$ in $\mathbb{R}^n$ and $\mathbf{c} \in U$ is either a local maximum or minimum of $f$ that occurs on the interior of $U$, then $\nabla f(\mathbf{c}) = \mathbf{0}$.

This theorem describes necessary conditions, not sufficient conditions. For example, consider the functions $f_1(x, y) = x^2 + y^2$ and $f_2(x, y) = y^2 - x^2$, wherein

$$\nabla f_1(x, y) = (2x, 2y) \quad \text{and} \quad \nabla f_2(x, y) = (-2x, -2y).$$

Both functions have critical points at $(x, y) = (0, 0)$, however the former is a minimum while the later is not. In particular, the latter function gives an example of a *saddle point.* Saddle points are those critical points which are neither maxima nor minima.

There is one additional kind of critical point which can appear. The above discussion of maxima, minima, and saddle points amounted to the function looking as though it had either a maximum or a minimum in every direction, and whether or not those directions all agreed with one another. This has not yet captured the idea of an inflection point.

**Definition 11.19**

If $f : \mathbb{R}^n \to \mathbb{R}$ is $C^2$ and $\mathbf{c}$ is a critical point of $f$, then we say that $\mathbf{c}$ is a *degenerate critical point* if $\mathbf{f}$ is rank $H_f(\mathbf{c}) < n$.

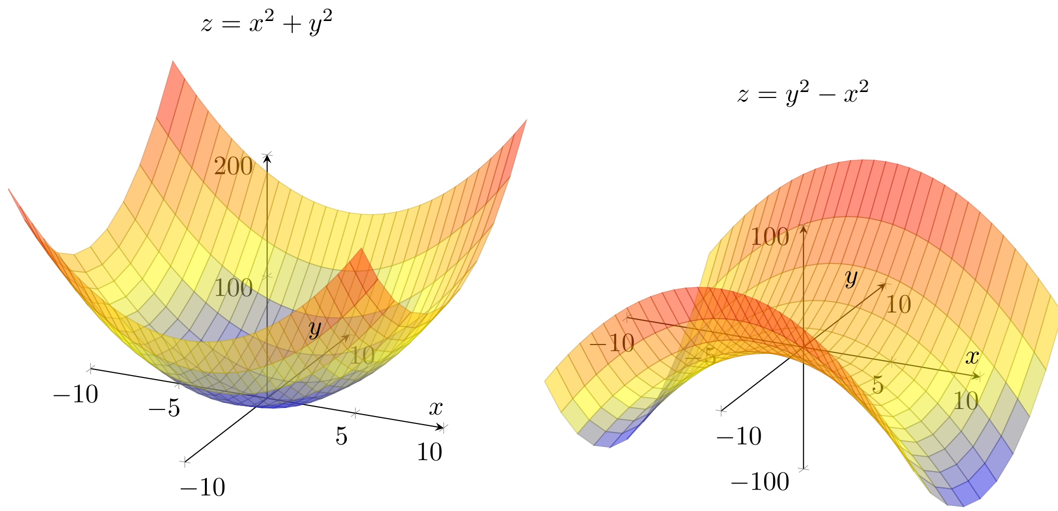



Figure 11.2: The graphs of $f_1(x, y) = x^2 + y^2$ and $f_2(x, y) = y^2 - x^2$. Both functions have a critical point at $(x, y) = (0, 0)$, but one is a minimum and the other is a saddle point.

Graphing functions is a terrible way to determine maxima and minima though, so we need to develop another criteria for determining extrema. This comes in the form of the second derivative test.

**Proposition 11.20**

Let $f$ be a function on $\mathbb{R}^n$ and $\mathbf{c}$ be a critical point. Let $H_f$ be the Hessian of $f$.

1. If the eigenvalues of $H_f(\mathbf{c})$ are all strictly positive, then $\mathbf{c}$ is a minimum.
2. If the eigenvalues of $H_f(\mathbf{c})$ are all strictly negative, then $\mathbf{c}$ is a maximum.
3. If the eigenvalues of $H_f(\mathbf{c})$ are a mix of positive and negative, then $\mathbf{c}$ is a saddle point.
4. If any of the eigenvalues of $H_f(\mathbf{c})$ are 0, the test is inconclusive.

While this can be labour intensive to check, there is a shortcut when $f$ is a function of two variables. The determinant of a matrix is equal to the product of its eigenvalues. Since a $2 \times 2$ matrix only has two eigenvalues, a negative determinant means that the eigenvalues must be of a different sign. If the eigenvalues are positive, they are the same sign, but one additional criteria must be checked to see if they are both negative or positive.

**Proposition 11.21**

Let $f$ be a function on $\mathbb{R}^2$ and $\mathbf{c}$ be a critical point. Define

$$D(x,y) = \det(H_f(x,y)) = \partial_{11}f(x,y)\partial_{22}f(x,y) - [\partial_{12}f(x,y)]^2,$$

be the *determinant* of $H_f$.

1. If $D(\mathbf{c}) < 0$ then $\mathbf{c}$ is a saddle point
2. If $D(\mathbf{c}) > 0$ then
   (a) If $\partial_{11}f(\mathbf{c}) > 0$ then $\mathbf{c}$ is a minimum,
   (b) If $\partial_{11}f(\mathbf{c}) < 0$ then $\mathbf{c}$ is a maximum.

If $\det H(\mathbf{c}) = 0$ then the result is inconclusive.

**Example 11.22**

Determine the critical points of the function $f(x,y) = x^4 - 2x^2 + y^3 - 6y$ and classify each as a maxima, minima, or saddle point.

*Solution.* The gradient can be quickly computed to be $\nabla f(x,y) = (4x(x^2-1), 3(y^2-2))$. The first component is zero when $x = 0, \pm 1$ and the second component is zero when $y = \pm\sqrt{2}$, giving six critical points: $(0, \pm\sqrt{2}), (-1, \pm\sqrt{2})$, and $(1, \pm\sqrt{2})$. The Hessian is easily computed to be

$$H_f(x,y) = \begin{bmatrix} 12x^2 - 4 & 0 \\ 0 & 6y \end{bmatrix} \quad \text{with} \quad D(x,y) = 6y(12x^2-4).$$

Evaluating $D(x,y)$ at our critical points gives

$$D(0,\pm\sqrt{2}) = \mp 24\sqrt{2}, \quad D(-1,\pm\sqrt{2}) = \pm 48\sqrt{2}, \quad D(1,\pm\sqrt{2}) = \pm 48\sqrt{2}.$$

Thus the maximum is $(0, -\sqrt{2})$, the minima are $(\pm 1, \sqrt{2})$, and the other three points are saddles. ■

**Example 11.23**

Show that the function $f(x,y) = y^2 - x^3$ has a degenerate critical point at $(x,y) = (0,0)$.

*Solution.* The gradient is $\nabla f(x,y) = (-3x^2, 2y)$ which indeed has a critical point at $(0,0)$. Furthermore, the Hessian is

$$H_f(x,y) = \begin{bmatrix} -6x & 0 \\ 0 & 2 \end{bmatrix}, \quad \text{so} \quad H_f(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

showing that $H(0,0)$ has rank 1. We conclude that $(0,0)$ is a degenerate critical point. ■

### 11.4.2 Constrained Optimization

The previous section introduced the notion of critical points, which can be used to determine maxima/minima on the interior of a set. However, what happens when we are given a set with empty interior? Similarly, if one is told to optimize over a set, it is not sufficient to only optimize over the interior, one must also check the boundary.

We have seen problems of constrained optimization before. Recall Example 7.36, which amounted to the following

> "You are building a fenced rectangular pasture, with one edge located along a river. Given that you have 200m of fencing, find the dimensions which maximize the volume of the pasture."

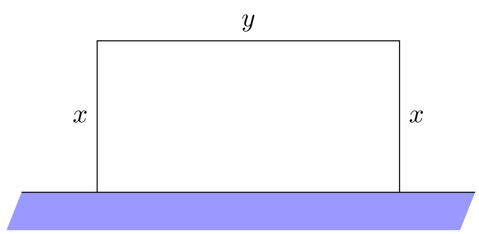



Figure 11.3: A visualization of simple optimization problem.

Translating this problem into mathematics, we let $x$ be the length and $y$ be the width of the pasture. We must then maximize the function $f(x, y) = xy$ subject to the constraint $2x + y = 200$. The equation $2x+y = 200$ is a line in $\mathbb{R}^2$, so we are being asked to determine the maximum value of the function $f$ along this line. The way that this was handled was to use the constraint to rewrite one variable in terms of another, and use this to reduce our function to a single variable. For example, if we write $y = 200 - 2x$ then

$$f(x, y) = x(200 - 2x) = 200x - 2x^2.$$

The lone critical point of this function occurs at $x = 50$, which gives a value of $y = 100$, and one can quickly check that this is the max.

Another technique that one could employ is the following: Recognizing that $2x + y = 200$ is just a line in $\mathbb{R}^2$, we can parameterize that line by a function $\gamma(t) = (t, 200 - 2t)$. The composition $f \circ \gamma$ is now a function in terms of the independent parameter $t$, yielding $f(\gamma(t)) = 200t - 2t^2$ which of course gives the same answer.

The fact that our constraint was just a simple line made this problem exceptionally simple. What if we wanted to optimize over a more difficult one-dimensional space, or even a two dimensional surface? Once again we can try to emulate the procedures above, and we may even meet with some success. However, there is a more novel way of approaching such problems, using the method of *Lagrange multipliers*.

**Theorem 11.24: Lagrange Multipliers**

Let $f, G$ be differentiable functions in $\mathbb{R}^n$, and let $S$ be the set of points $\mathbf{x}$ such that $G(\mathbf{x}) = 0$. If $f$, when restricted to $S$, has a maximum or minimum at the point $\mathbf{c} \in S$ and $\nabla G(\mathbf{c}) \neq 0$, then there exists $\lambda \in \mathbb{R}$ such that

$$\nabla f(\mathbf{c}) = \lambda \nabla G(\mathbf{c}).$$

Okay, what does that all mean? Here the function $G$ describes the constraint set. In the example above, we knew that $x$ and $y$ had to satisfy $2x + y = 200$. Setting $G(x, y) = 2x + y - 200$, then $G(x, y) = 0$ is the same thing as $2x + y = 200$. The equation $\nabla f(\mathbf{c}) = \lambda \nabla G(\mathbf{c})$ gives the extra equation which we hope will let us find the points where the max and min occur.

**Example 11.25**

Use the method of Lagrange multipliers to solve the problem given in Figure 11.3.

*Solution.* The constraint in our fencing problem is given by the function $G(x, y) = 2x + y - 200 = 0$. We can easily compute $\nabla f(x, y) = (y, x)$ and $\nabla G(x, y) = (2, 1)$, so by the method of Lagrange multipliers, there exists $\lambda \in \mathbb{R}$ such that $\nabla f(x, y) = \lambda \nabla G(x, y)$; that is,

$$\begin{bmatrix} y \\ x \end{bmatrix} = \lambda \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

We thus know that $y = 2\lambda, x = \lambda$, and substituting this into $2x - y = 200$ gives $4\lambda = 200$. Thus $\lambda = 50$, from which we conclude that $y = 2\lambda = 100$ and $x = \lambda = 50$ as required. ■

**Example 11.26**

Maximize the function $f(x, y, z) = xyz$ on the ellipsoid $x^2 + 2y^2 + 3z^2 = 1$.

*Solution.* The constraint equation is given by $G(x, y, z) = x^2 + 2y^2 + 3z^2 - 1 = 0$. When we compute our gradients, the method of Lagrange multipliers gives the following system of equations:

$$\begin{aligned} yz &= 2\lambda x \\ xz &= 4\lambda y \\ xy &= 6\lambda z \end{aligned}$$

If we combine this with the constraint $x^2 + 2y^2 + 3z^2 = 1$ we have four equations in four unknowns, though all the equations are certainly non-linear! Herein we must be clever, and start manipulating our equations to try and solve for $(x, y, z)$. Notice that if we play with the term $xyz$ then depending on how we use the associativity of multiplication, we can get an additional set of conditions. For example

$$\begin{aligned} x(yz) &= x(2\lambda x) = 2\lambda x^2 \\ y(xz) &= y(4\lambda y) = 4\lambda y^2 \\ z(xy) &= z(6\lambda z) = 6\lambda z^2 \end{aligned}$$

and all of these must be equal. We can make a small simplification by removing a factor of 2 to get

$$\lambda x^2 = 2\lambda y^2 = 3\lambda z^2. \tag{11.2}$$

**Case 1** ($\lambda = 0$): If $\lambda = 0$ then $yz = xz = xy = 0$. This immediately implies that two of $x, y$, or $z$ must be zero, so $f(x, y, z) = xyz = 0$. If $x = y = 0$ then the constraint equation gives $\left(0, 0, \pm\frac{1}{\sqrt{3}}\right)$. If $x = z = 0$ then $\left(0, \pm\frac{1}{\sqrt{2}}, 0\right)$ and if $y = z = 0$ then $(\pm 1, 0, 0)$. So all of these points give a result of $f(x, y, z) = 0$ and are candidates for maxima/minima.

**Case 2** ($\lambda \neq 0$): If $\lambda \neq 0$ then we can divide (11.2) by $\lambda$ to get that $x^2 = 2y^2 = 3z^2$. Substituting this into the constraint equation we get $1 = x^2 + x^2 + x^2 = 3x^2$ so that $x = \pm\frac{1}{\sqrt{3}}$, which we can use to find $y$ and $z$. This gives us eight possible critical points corresponding to the following choice of signs:

$$x = \pm\frac{1}{\sqrt{3}}, \quad y = \pm\frac{1}{\sqrt{6}}, \quad z = \pm\frac{1}{3}.$$

There are only two possible values of $f$ for these points, namely $f(x, y, z) = \pm\frac{1}{9\sqrt{2}}$. Since these are both either bigger than 0 or smaller than 0, these are the corresponding global maxima/minima of the function. ■

**Example 11.27**

Determine the maximum and minimum of the function $f(x, y) = x^2 + 2y^2$ on the disk $x^2 + y^2 \leq 4$.

*Solution.* We begin by determining critical points on the interior. Here we have $\nabla f(x, y) = (2x, 4y)$ which can only be $(0, 0)$ if $x = y = 0$. Here we have $f(0, 0) = 0$.

Next we determine the extreme points on the boundary $x^2 + y^2 = 4$, for which we set up the constraint function $G(x, y) = x^2 + y^2 - 4$ with gradient $\nabla G(x, y) = (2x, 2y)$. Using the method of Lagrange multipliers, we thus have

$$\begin{aligned} 2x &= 2\lambda x \\ 4y &= 2\lambda y \end{aligned}$$

**Case 1** ($x \neq 0$): If $x \neq 0$ then we can solve $2x = 2\lambda x$ to find that $\lambda = 1$. This implies that $y = 2y$ which is only possible if $y = 0$. Plugging this into the constraint gives $x^2 = 4$ so that $x = \pm 2$, so our candidate points are $(\pm 2, 0)$, which give values $f(\pm 2, 0) = 4$.

**Case 2** ($y \neq 0$): If $y \neq 0$ then we can solve $4y = 2\lambda y$ to find that $\lambda = 2$. This implies that $2x = 4x$ which is only possible if $x = 0$. Solving the constraint equation thus gives the candidates $(0, \pm 2)$, which gives values $f(0, \pm 2) = 8$.

The case where $\lambda = 0$ gives no additional information. Hence we conclude that the minimum occurs at $(0, 0)$ with a value of $f(0, 0) = 0$, while the maximum occurs at the two points $(0, \pm 2)$ with a value of $f(0, \pm 2) = 8$. ■

If multiple constraints are given, the procedure is similar, except that we now need additional multipliers. Suppose $G_1, \ldots, G_m$ are all functions of $n$ variables and set

$$S = \{\mathbf{x} \in \mathbb{R}^n : G_1(\mathbf{x}) = G_2(\mathbf{x}) = \cdots = G_m(\mathbf{x}) = 0\}.$$

If we are tasked with optimizing $f$ on $S$, then Lagrange Multipliers implies that $\mathbf{c} \in S$ is a maximum or minimum there exist $\lambda_1, \ldots \lambda_m \in \mathbb{R}$ such that

$$\nabla f(\mathbf{c}) = \sum_{i=1}^{m} \lambda_i \nabla G_i(\mathbf{c}).$$

**Example 11.28**

Find the maximum of $f(x, y, z) = xyz$ subject to the constraint $y = 2x$ and $x + y + z = 45$.

*Solution.* Set $G_1(x, y, z) = x + y + z - 45$ and $G_2(x, y, z) = y - 2x$ so that

$$\nabla f(x, y, z) = (yz, xz, xy), \quad \nabla G_1(x, y, z) = (1, 1, 1), \quad \nabla G_2(x, y, z) = (-2, 1, 0.$$

By the theory of Lagrange multipliers, there exist $\lambda_1$ and $\lambda_2$ such that $\nabla f(x, y, z) = \lambda_1 \nabla G_1(x, y, z) + \lambda_2 \nabla G_2(x, y, z)$, or equivalently

$$yz = \lambda_1 - 2\lambda_2, \quad xz = \lambda_1 + \lambda_2, \quad xy = \lambda_1.$$

Subbing the third equation into the second gives $xz = xy + \lambda_2$ or $\lambda_2 = xz - xy$. Subbing this into the first equation gives

$$yz = xy - 2(xz - xy) = 3xy - 2xz.$$

Now we use the fact that $y = 2x$ to remove all dependency on $y$:

$$2xz = 6x^2 - 2xz \quad \Rightarrow \quad 4xz = 6x^2.$$

If $x = 0$ then $f(x, y, z) = 0$, so we keep this in mind. Assuming that $x \neq 0$, we can cancel an $x$ from each side to get $2z = 3x$ or $z = 3x/2$. Subbing $y = 2x$ and $z = 3x/2$ into $x + y + z = 45$ we get

$$45 = x + y + z = x + 2x + \frac{3}{2}x = \frac{9}{2}x$$

which means $x = 10$. This in turn implies that $y = 2x = 20$ and $z = 3x/2 = 15$. This solution $(10, 20, 15)$ returns $f(10, 20, 15) = 3000$, which is the maximum value of the function. ∎

## 11.5 Iterated Integrals

Iterated integrals is the process of evaluating a multidimensional integral. This section is about as haphazardly done as possible, since determining when an integral can be broken into a bunch of one dimensional integrals is a subtle and complicated situation.

Anyway, the idea here is that integration reverses the process of differentiation. In learning about partial derivatives, we held all other variables constant and just differentiated along a single direction. Similarly, with iterated integrals we hold all variables constant save for one, then integrate. On a rectangle, this isn't too bad.

**Example 11.29**

Determine the volume under the function $f(x,y) = xe^{x^2-y}$ on the rectangle $R = [0,1]\times[0,1]$.

*Solution.* Since $f$ is a continuous function on $R$ it is integrable, and so certainly each of the slices $f_y$ or $f_x$ are integrable as well. We will do the calculation both ways to show that the integral yields the same results. If we integrate first with respect to $x$ then $y$, we have

$$\begin{aligned}\int_0^1 \left[\int_0^1 xe^{x^2-y}\,\mathrm{d}x\right]\mathrm{d}y &= \int_0^1 \left[\frac{1}{2}e^{x^2-y}\right]_{x=0}^1 \mathrm{d}y \\ &= \frac{1}{2}(e-1)\int_0^1 e^{-y}\,\mathrm{d}y \\ &= \frac{1}{2}(e-1)\left[-e^{-y}\right]_0^1 = -\frac{1}{2}(e-1)(e^{-1}-1)\end{aligned}$$

Conversely, let us instead integrate with respect to $y$ first. We have

$$\begin{aligned}\int_0^1 \left[\int_0^1 xe^{x^2-y}\,\mathrm{d}y\right]\mathrm{d}x &= -(e^{-1}-1)\int_0^1 xe^{x^2}\,\mathrm{d}x \\ &= -\frac{1}{2}(e^{-1}-1)(e-1)\end{aligned}$$

As expected, the result was the same either way. ■

Of course, the above example was very simple since we could decompose our function $f(x,y) = f_1(x)f_2(y)$, but the result still holds even when such a decomposition is not possible.

Now rectangles are rather boring objects about which to integrate, so we again look at other sets $S \subseteq \mathbb{R}^2$. In particular, we will suppose that $S$ has a relatively nice boundary, defined as

$$S = \{(x,y) : a \leq x \leq b, \alpha(x) \leq y \leq \beta(x)\}.$$

In this case, our integration becomes

$$\int_S f\,\mathrm{d}A = \int_a^b \left[\int_{\alpha(x)}^{\beta(x)} f(x,y)\,\mathrm{d}y\right]\mathrm{d}x.$$

Often times, the most difficult part of solving an iterated integral question comes from determining the bounding functions, though sometimes we are fortunate and they are already prescribed.

**Example 11.30**

Find the integral of the function $f(x,y) = \dfrac{y}{x^5+1}$ on the region bounded by the lines $y=0$ $x=1$ and $y=x^2$.

*Solution.* In any situation of performing iterated integrals, it is best to draw a diagram of the region over which we are integrating. In our case, we can see that the region may be summarily described as

$$S = \left\{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq x^2\right\}.$$

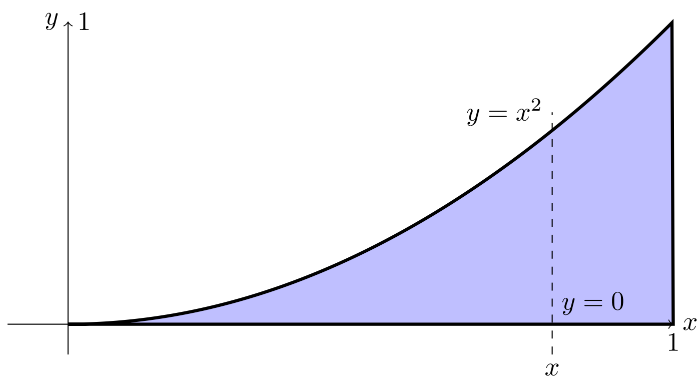



Figure 11.4: Determining the bounded functions by drawing a line. Where does the line enter and leave the corresponding shape?

Certainly our function is continuous on $S$ (since $x^5 + 1 \neq 0$ on this set) and so is integrable, along with any of the slices. Integrating gives

$$\begin{aligned}
\iint_S f \, \mathrm{d}A &= \int_0^1 \left[\int_0^{x^2} \frac{y}{x^5+1} \, \mathrm{d}y\right] \mathrm{d}x \\
&= \frac{1}{2} \int_0^1 \left[\frac{y^2}{x^5+1}\right]_0^{x^2} \mathrm{d}x = \frac{1}{2} \int_0^1 \frac{x^4}{x^5+1} \, \mathrm{d}x \\
&= \frac{1}{10} \ln|x^5+1| \Big|_0^1 = \frac{\ln(2)}{10}.
\end{aligned}$$

■

Note that the region in Example 11.30 also could have been described by

$$S = \{(x, y) : 0 \leq y \leq 1, \ \sqrt{y} \leq x \leq 1\},$$

so we also could have (attempted to) compute the integral as

$$\iint_S f \, \mathrm{d}A = \int_0^1 \left[\int_{\sqrt{y}}^1 \frac{y}{x^5+1} \, \mathrm{d}x\right] \mathrm{d}y.$$

This probably would not have worked as nicely though, since $1/(x^5 + 1)$ is not easy to integrate. This suggests that being able to rewrite our domain is a useful skill, since sometimes we are given the boundary, but the problem is not amenable to the given description.

**Example 11.31**

Determine the integral of the function $f(x, y) = e^{y^2}$ on the region bounded by the lines $y = 1$, $x = 0$ and $y = x$.

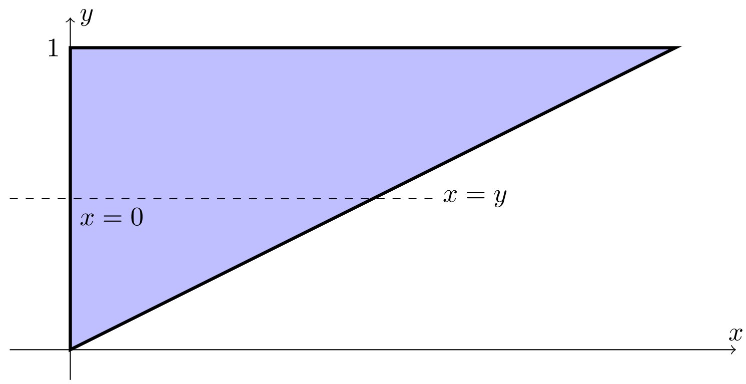



Figure 11.5: Sometimes it's more convenient to integrate in the $y$-direction first.

*Solution.* The region is a simple triangle, given in Figure 11.5, which can be written as either of the following two sets

$$S = \{(x, y) : 0 \leq x \leq 1,\ x \leq y \leq 1\} = \{(x, y) : 0 \leq y \leq 1,\ 0 \leq x \leq y\}.$$

If we try to use the first description, we get

$$\int_S f \, \mathrm{d}A = \int_0^1 \left[ \int_x^1 e^{y^2} \, \mathrm{d}y \right] \mathrm{d}x$$

but the function $e^{y^2}$ has no elementary anti-derivative, and we are stuck. On the other hand, using the second description gives

$$\begin{aligned} \int_S f \, \mathrm{d}A &= \int_0^1 \left[ \int_0^y e^{y^2} \, \mathrm{d}x \right] \mathrm{d}y \\ &= \int_0^1 \left[ x e^{y^2} \right]_{x=0}^{x=y} \mathrm{d}y = \int_0^1 y e^{y^2} \, \mathrm{d}y \\ &= \left[ \frac{1}{2} e^{y^2} \right]_{y=0}^1 = \frac{1}{2}(e - 1). \end{aligned}$$

■

**Example 11.32**

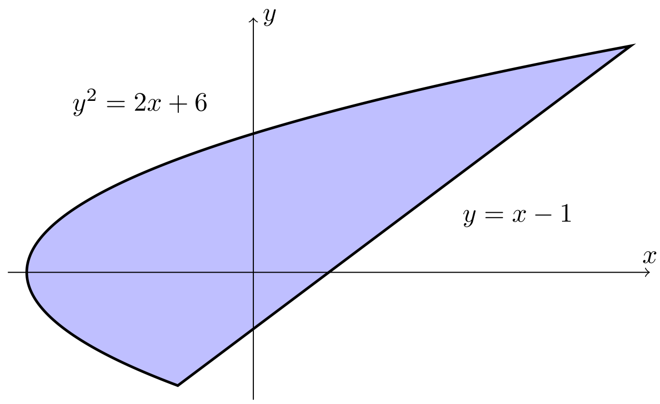
Determine $\iint_S xy \, \mathrm{d}A$ where $S$ is the region bounded by $y = x - 1$ and $y^2 = 2x + 6$.

*Solution.* We begin by drawing a rough picture of what the boundary looks like. Notice that the intersection of these two lines occurs when

$$(x-1)^2 = 2x + 6, \qquad \Leftrightarrow \qquad x^2 - 4x - 5 = 0, \qquad \Leftrightarrow \qquad x = 5, -1,$$

which corresponds to the pairs $(-1, -2)$ and $(5, 4)$. Now our figure shows that it will be very hard to write this as $\{a \leq x \leq b, \alpha(x) \leq y \leq \beta(x)\}$, so instead we try to switch the variables. In that case, notice that we can write $S$ as

$$S = \left\{ (x, y) : -2 \leq y \leq 4,\ \frac{1}{2} y^2 - 3 \leq x \leq y + 1 \right\}.$$




Now integrating, we get

$$\begin{aligned}
\int_S xy\,\mathrm{d}A &= \int_{-2}^{4}\left[\int_{\frac{1}{2}y^2-3}^{y+1} xy\,\mathrm{d}x\right]\mathrm{d}y \\
&= \frac{1}{2}\int_{-2}^{4}\left[x^2y\right]_{\frac{1}{2}y^2-3}^{y+1}\mathrm{d}y \\
&= \frac{1}{2}\int_{-2}^{4} y\left[(y+1)^2 - \left(\frac{1}{2}y^2-3\right)^2\right]\mathrm{d}y \\
&= \frac{1}{2}\int_{-2}^{4}\left[-\frac{y^5}{4} + 4y^3 + 2y^2 - 8y\right]\mathrm{d}y \\
&= \frac{1}{2}\left[-\frac{y^6}{24} + y^4 + \frac{2y^3}{3} - 4y^2\right]_{-2}^{4} = 36.
\end{aligned}$$

■

**Triple! Integrals:** Of course we have limited our discussion thus far to functions of two variables, but there was no reason to (other than to keep ourselves from headaches). Naturally, we can extend to three dimensions and beyond, and so perform integration in $n$-variables. However, because drawing diagrams is so critical for doing iterated integrals, we typically tend to avoid doing them in 4-dimensions or greater. In this course, we will not see integrals in more than 3-variables.

This being said, what happens when we want to integrate a function in three variables? The solution is to proceed just as before, except that now we write our domain as

$$S = \{(x,y,z) : a \leq x \leq b,\ \alpha(x) \leq y \leq \beta(x),\ \varphi(x,y) \leq z \leq \psi(x,y)\},$$

and the corresponding integral becomes

$$\iiint_S f(x,y,z)\,\mathrm{d}A = \int_a^b \left[\int_{\alpha(x)}^{\beta(x)}\left[\int_{\varphi(x,y)}^{\psi(x,y)} f(x,y,z)\,\mathrm{d}z\right]\mathrm{d}y\right]\mathrm{d}x.$$

**Example 11.33**

Determine $\iiint_S z\ \mathrm{d}A$ if $S$ is the set bounded by the planes $x = 0, y = 0, z = 0$ and $x+y+z = 1$.

*Solution.* This shape is a tetrahedron whose boundaries are the three standard unit normals $\{e_i\}_{i=1,2,3}$ and the origin $(0,0,0)$. Now $0 \leq x \leq 1$ is evident, and projecting into the $xy$-plane we see that $0 \leq y \leq 1-x$. Finally, we clearly have that $0 \leq z \leq 1-x-y$ so that

$$\begin{aligned}
\iiint_S z \ \mathrm{d}A &= \int_0^1 \left[\int_0^{1-x}\left[\int_0^{1-x-y} z\,\mathrm{d}z\right]\mathrm{d}y\right]\mathrm{d}x \\
&= \int_0^1 \left[\int_0^{1-x}\left[\frac{z^2}{2}\right]_0^{1-x-y}\mathrm{d}y\right]\mathrm{d}x \\
&= \frac{1}{2}\int_0^1\left[\int_0^{1-x}(1-x-y)^2\,\mathrm{d}y\right]\mathrm{d}x = \frac{1}{2}\int_0^1\left[-\frac{(1-x-y)^3}{3}\right]_0^{1-x}\mathrm{d}x \\
&= \frac{1}{6}\int_0^1 (1-x)^3\,\mathrm{d}x = \frac{1}{6}\left[-\frac{(1-x)^4}{4}\right]_0^1 = \frac{1}{24}.
\end{aligned}$$

■

**Example 11.34**

Determine $\iiint_S (2x+4z)\,\mathrm{d}V$ where $S$ is the region bounded by the planes $y=x$, $z=x$, $z=0$, and $y=x^2$.

*Solution.* You should stare at these equations for some time and try to visualize the space. In particular, a nice parameterization of the space can be given as

$$S = \left\{(x,y) : 0 \leq x \leq 1,\ x^2 \leq y \leq x,\ 0 \leq z \leq x\right\}.$$

Integrating gives

$$\begin{aligned}
\iiint_S f\,\mathrm{d}V &= \int_0^1\left[\int_{x^2}^x\left[\int_0^x (2x+4z)\,\mathrm{d}z\right]\mathrm{d}y\right]\mathrm{d}x \\
&= \int_0^1\left[\int_{x^2}^x 2x^2+2x^2\,\mathrm{d}y\right]\mathrm{d}x \\
&= 2\int_0^1\left(4x^3-4x^4\right)\mathrm{d}x \\
&= 4\left[\frac{1}{4}x^4 - \frac{1}{5}x^5\right]_0^1 = \frac{1}{5}.
\end{aligned}$$

■

## 11.6 Change of Variables

### 11.6.1 Coordinates

It is difficult to describe what we mean by a set of coordinates without using more technical language. The effective idea is that a coordinate system should be a way of (uniquely) and continuously describing a point in your space. Cartesian coordinates are those with which we are most familiar, and are given by $(x,y)$, describing the horizontal and vertical displacement of a point from the origin. However, the origin itself corresponds to an arbitrary choice: choose some other point in the plane and call that the origin, and notice that fundamentally, our space has not changed. For example, a circle $x^2 + y^2 = 1$ is in many ways the same as the circle $(x-a)^2 + (y-b)^2 = 1$ for any choice of $(a,b)$, we have simply "moved it." Such a transformation is called a *translation* and are described as functions $f(x,y) = (x-a, y-b)$.

Similarly, one might choose to change how we want to measure distances, resulting in a *scaling* of the from $f(x,y) = (\alpha x, \beta y)$ for $\alpha, \beta \neq 0$ (when $\alpha < 0$ this corresponds to reflecting about the $y$-axis, and similarly $\beta < 0$ is reflection about the $x$-axis). We could even rotate our coordinate system by an angle $\theta$, though doing so requires trigonometry. Combining scaling, rotations, and translations, one gets *affine transformations* $f(x,y) = (c_1 x + c_2 y + c_3, d_1 x + d_2 y + d_3)$.

There are countless other types of coordinate systems one might want to use, for example $(u,v) = G(x,y) = (e^x, y^2)$, though we run into uniqueness issues and need to restrict our sets in order to have a "good" coordinate system. In this case, a good coordinate system is between the sets $\mathbb{R} \times [0,\infty)$ and $(0,\infty) \times [0,\infty)$.

We need to determine how areas change under these new coordinate systems. This is done as follows:

**Definition 11.35**

Suppose $u = f(x,y)$ and $v = g(x,y)$ is a coordinate transformation. The *Jacobian matrix* of this transformation is
$$M(x,y) = \begin{bmatrix} f_x(x,y) & f_y(x,y) \\ g_x(x,y) & g_y(x,y) \end{bmatrix},$$
while the *Jacobian* is $J(x,y) = f_x(x,y)g_y(x,y) - f_y(x,y)g_x(x,y)$.

In the case of a scaling transformation $(u,v) = (\alpha x, \beta y)$, note that
$$M(x,y) = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}, \quad J(x,y) = \alpha\beta.$$
For the translation transformation $(u,v) = (x+\alpha, y+\beta)$ we get
$$M(x,y) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad J(x,y) = 1.$$
The Jacobian measures the infinitesimal change in area; namely $\mathrm{d}u\,\mathrm{d}v = |J(x,y)|\,\mathrm{d}x\,\mathrm{d}y$.

**Example 11.36**

Consider the coordinate transformation $(u,v) = (x\ln(y), xy^2)$. Find the corresponding Jacobian.

*Solution.* The Jacobian matrix is
$$M(x,y) = \begin{bmatrix} \ln(y) & x/y \\ y^2 & 2xy \end{bmatrix},$$
so that $J(x,y) = 2xy\ln(y) - xy = xy(2\ln(y) - 1)$. ■

### 11.6.2 Integration

Let's motivate the situation by analyzing what happens in the one-dimensional case. In the one-dimensional case, there is not much in the way of variable changing that can be done! Nonetheless, you have already seen a plethora of examples which emulate coordinate changing: The method of substitution. For example, when integrating the equation

$$\int_2^3 \frac{x}{x^2-1}\,\mathrm{d}x,$$

you should (hopefully) realize that the appropriate substitution here is $u = x^2 - 1$ so that $\mathrm{d}u = 2x\,\mathrm{d}x$, and the integral becomes

$$\int_2^3 \frac{x}{1-x^2}\,\mathrm{d}x = \frac{1}{2}\int_3^8 \frac{1}{u}\,\mathrm{d}u = [\ln|u|]_3^8 = \ln(8) - \ln(3).$$

In effect, we have realized that working in the $x$-coordinate system is silly since it makes our integral look complicated. By changing to the $u = 1 + x^2$ coordinate system, the integral reduces to something which we can easily solve.

The theory is as follows (though our presentation might seem a bit backwards compared to how such integrals are usually computed): The fundamental theorem of calculus tells us that

$$\int_a^b f(g(x))g'(x)\,\mathrm{d}x = \int_{g(a)}^{g(b)} f(u)\,\mathrm{d}u \tag{11.3}$$

where $u = g(x)$ so that $\mathrm{d}u = g'(x)\,\mathrm{d}x$. The idea is that by introducing the auxiliary function $u = g(x)$ we were able to reduce the problem to something more elementary, and that is the goal of changing variables.

Unfortunately, there is never a single way to change variables, and it can make our notation a bit of a headache. For example, what if we had instead chosen the substitution $u = 1 - x^2$ in the previous example, so that the integral became

$$\int_2^3 \frac{x}{x^2-1}\,\mathrm{d}x = \frac{1}{2}\int_{-3}^{-8} \frac{1}{u}\,\mathrm{d}u.$$

Notice that the bounds of integration are in the wrong order, since certainly $-3 > -8$. We fix this by introducing a negative sign and interchanging the bounds and arrive at the same answer, but the point is that we do not want to have to worry about whether we have changed the *orientation*[5] of the interval (since this will become a grand nightmare in multiple dimensions!). Hence if $I = [a, b]$, we will write (11.3) as

$$\int_I f'(g(x))|g'(x)|\,\mathrm{d}x = \int_{g(I)} f(u)\,\mathrm{d}u.$$

What is bothersome about this equation is that $g$ appears on both sides of the equation. If $g$ is a change of coordinates, then there is no harm in replacing $g$ with $g^{-1}$. Let $J = g(I)$ so that we get

$$\int_{g^{-1}(J)} f'(g(x))|g'(x)|\,\mathrm{d}x = \int_J f(u)\,\mathrm{d}u.$$

[5]This is a remarkably subtle but important point that does not manifest in 1-dimension but proves to be truly inconvenient in higher dimensions. There is an entire theory of *orientability* of surfaces and higher dimensional spaces, and if your space is not orientable then it is difficult to do integration.

So what do we do in higher dimension?

**Theorem 11.37: Change of Variables**

If $S, T \subseteq \mathbb{R}^2$ and $(u, v) = G(x, y)$ is a change of variables going from $S$ to $T$, then for any integrable function $f(x, y)$ on $T$ we have

$$\int_S f(g(x,y))J(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{G(S)} f(u,v)\,\mathrm{d}u\,\mathrm{d}v$$

where $G(S)$ is what $S$ becomes under the function $G$.

**Example 11.38**

Let $S$ be the region bounded by the curves $xy = 1$, $xy = 3$, $x^2 - y^2 = 1$ and $x^2 - y^2 = 4$. Compute $\iint_T (x^2 + y^2)\,\mathrm{d}A$.

*Solution.* The region suggests that we should take a change of variables of the form $u = xy$ and $v = x^2 - y^2$, so that setting

$$T = \{1 \leq u \leq 3,\ 1 \leq v \leq 4\}$$

implies that $G : S \to T$ given by $(u, v) = G(x, y) = (xy, x^2 - y^2)$ is the change of variables we want. Now

$$J(x,y) = 2(x^2 + y^2).$$

Thus $\mathrm{d}u\,\mathrm{d}v = 2(x^2 + y^2)\,\mathrm{d}x\,\mathrm{d}y$ and our integral becomes

$$\iint_S (x^2 + y^2)\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{2}\int_T \mathrm{d}u\ \mathrm{d}v = 3. \qquad \blacksquare$$

**Example 11.39**

Determine $\iint_S \frac{(x+y)^4}{(x-y)^5}\,\mathrm{d}x\,\mathrm{d}y$ if $S = \{-1 \leq x + y \leq 1, 1 \leq x - y \leq 3\}$.

*Solution.* Let $u = x + y$ and $v = x - y$ so that $S$ becomes the rectangle $\{-1 \leq u \leq 1, 1 \leq v \leq 3\}$. The Jacobian is $J(x, y) = -2$, so that under a change of variable we get

$$\iint_S \frac{(x+y)^4}{(x-y)^5}\,\mathrm{d}x\,\mathrm{d}y = \int_{-1}^{1}\int_1^3 \frac{2u^4}{v^5}\,\mathrm{d}u\,\mathrm{d}v = 2\int_{-1}^{1} u^4\,\mathrm{d}u \int_1^3 \frac{1}{v^5}\,\mathrm{d}v = \frac{16}{81}. \qquad \blacksquare$$

## 11.7 Exercises

11-1. Consider the function $f(x, y, z) = x^2 - y^2 - z$. For $k = 0, 1, 2, 3, 4$ plot

(a) $f(x,y,k)=0$
(b) $f(x,k,z)=0$
(c) $f(k,y,z)=0$

Use your above results to plot $z=x^2-y^2$.

11-2. Use the limit definition of the partial derivative to find each derivative below.

(a) $\dfrac{\partial f}{\partial x}$ where $f(x,y,z)=x^2yz$
(b) $\dfrac{\partial f}{\partial y}$ where $f(x,y,z)=x^2+y^2+z^2$
(c) $\dfrac{\partial f}{\partial z}$ where $f(x,y,z)=\dfrac{xy}{z}$.

11-3. Compute the stated partial derivative in each case.

(a) $f_x$ where $f(x,y,z)=x^2y+y^2z+z^2x$.
(b) $g_u$ where $g(u,v)=uve^u$.
(c) $h_s$ where $h(s,t)=s^2t\ln(st)$.
(d) $\partial_x f$ where $f(x,y,z,w)=xye^{xz+wy}$.
(e) $\partial_v g$ where $g(u,v)=\sqrt{u^2+v^2+uv}$
(f) $\partial_t h$ where $h(s,t)=s^t$

11-4. Compute $\nabla f$ for each $f$ below.

(a) $f(x,y)=x^2y+y\ln(x)$
(b) $f(x,y,z)=xy+yz+zx$
(c) $f(x,y,z)=\ln(xyz)$

11-5. Given the following demand functions, determine if the products are complimentary or competitive.

(a) $q_A=30-p_A-10p_B;\quad q_B=100-3p_A-4p_B$
(b) $q_A=\frac{150}{p_A\sqrt{p_B}};\quad q_B=\frac{250}{p_B\sqrt[3]{p_A}}$

11-6. Let $\mathbf{x}\in\mathbb{R}^n$. A function $f$ of $n$-variables is said to be homogeneous of degree $k$ if $f(\lambda\mathbf{x})=\lambda^k f(\mathbf{x})$. For example, $f(x,y)=x^2y+y^2x$ is homogeneous of degree 3, since

$$f(\lambda x,\lambda y)=(\lambda x)^2(\lambda y)+(\lambda x)(\lambda y)^2=\lambda^3(x^2y+y^2x)$$

Which of the following functions are homogeneous? If they are homogeneous, of what degree are they?

(a) $f(x,y)=x^2+y^2$
(b) $f(x,y,z)=ye^{xz}$
(c) $f(x,y,z)=x^2y+y^2x+z^2$
(d) $f(x,y,z)=\ln(xyz)$
(e) $f(x,y,z,w)=x^2y^2+z^2w^2+xyzw$

11-7. Compute every second order derivative of the following functions:

(a) $f(x,y,z)=xyz$
(b) $f(x,y,z)=x^2+y^2+z^2$
(c) $f(x,y,z)=x\ln(yz)$
(d) $f(x,y,z)=ze^{xy}$

11-8. In each case, compute the prescribed higher order derivative:

(a) $f_{xyy}$ if $f(x,y) = x^2y^2 - 2xy$
(b) $f_{zzz}$ if $f(x,y,z) = xe^{z^2-y}$
(c) $f_{ust}$ if $f(u,v,s,t) = uv^2s^3t^4$
(d) $f_{rrss}$ if $f(r,s) = r^4\ln(s^2)$
(e) $f_{rssst}$ if $f(r,s,t) = \sqrt{r^2+t^2}e^s$

11-9. For each function $f$ below, write down the Hessian of $f$.

(a) $f(u,v) = u^2v - v^2u$
(b) $f(x,y,z) = xy + xz + yz$
(c) $f(a,b,c,d) = a^2 + b^2 + c^2 + d^2$

11-10. Find the critical points of the following functions:

(a) $g(x,y) = x^2 + y^2 - 16x - 14y + 31$
(b) $h(x,y) = x^3 - y^2 - xy + 13$
(c) $q(x,y,z) = x^2 + y^2 + 7x^2 - xy - 3yz$

11-11. Find and classify the critical points of the following functions:

(a) $h(x,y) = 100 - 3x^2 - 4y^2$
(b) $g(x,y) = 1/x + xy + 1/y$
(c) $f(x,y) = 9x^2y + 3y^3 - 9x^2 - 9y^2 - 17$

11-12. If $f$ is a function of two variables, its second order polynomial approximation at $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$ is

$$f(\mathbf{a}) + f_x(\mathbf{a})(x-a_1) + f_y(\mathbf{a})(y-a_2) + \frac{f_{xx}(\mathbf{a})}{2}(x-a_1)^2 + \frac{f_{yy}(\mathbf{a})}{2}(y-a_2)^2 + f_{xy}(\mathbf{a})(x-a_1)(y-a_2)$$

Find the second order polynomial approximation for the following functions:

(a) $f(x,y) = e^{xy}$ at $\mathbf{a} = (0,0)$
(b) $f(x,y) = x\ln(y)$ at $\mathbf{a} = (1,e)$

11-13. Determine the prescribed partial derivative using the chain rule.

(a) Find $\partial_s w$ if $w = xe^{xy}$ and $x = \sqrt{s^2+t}, y = s/t$.
(b) Find $\partial_u w$ if $w = xy\sqrt{x}$ and $x = \frac{t}{s+1}, y = st\, s = e^u, t = u^2$.
(c) Find $\partial_z w$ if $w = xy + yz$ and $x = y^2 + z^2, y = z$.

11-14. Suppose $w = f(x,y), x = g(t), y = h(x,t)$. Write down an expression for $\frac{\partial w}{\partial t}$.

11-15. Find and classify the critical points of $f(x,y) = x^2 + 3xy + y^2 - x + 3y$.

11-16. Suppose $f$ is a function of $n$-variables. How many $k$th-order derivatives does $f$ have?

11-17. Recall that an $n \times n$ matrix $A$ is said to be symmetric if $A = A^T$. Suppose $f$ is a $C^2$ function of $n$-variables. Argue that its Hessian matrix $H_f$ is symmetric.

11-18. Determine the maximum and minimum to the function $f(x,y) = x^2 + y^2 - 2x - 2y$ on the circle $x^2 + y^2 = 4$.

11-19. Find the maximum and minimum to the function $f(x, y) = xy$ if $2x^2 + 8y^2 = 16$

11-20. Find the point on the plane $x + 2y + 2z = 3$ which lies closest to the origin. *Hint:* The distance between $(x_0, y_0, z_0)$ and $(x_1, y_1, z_1)$ is $\sqrt{(x_0 - x_1)^2 + (y_0 - y_1)^2 + (z_0 - z_1)^2}$.

11-21. What is the volume of the largest box that can be inscribed in the sphere $x^2 + y^2 + z^2 = 1$?

11-22. Determine the global maximum and minimum of the function $f(x, y, z) = x^2 - y^2$ on the unit ball $x^2 + y^2 + z^2 \leq 1$

11-23. You're assembling the perfect trail mix and need to determine the optimal number of nuts to add. You have the option between Almonds, Brazil Nuts, and Cashews, which you buy in quantities $x, y, z$, at prices $p_x, p_y$, and $p_z$, respectively. Your nut budget is $M$ dollars.

(a) Determine an equation which describes your budget constraint.[6]

(b) Define an appropriate constraint function $g(x, y, z)$ so that your solution to part (11-23a) is written as $g(x, y, z) = M$. What is $\nabla g$?

(c) You *really* like Brazil nuts, so your happiness (utility) with the mix is dictated by $f(x, y, z) = xy^2z$. Write down $\nabla f$.

(d) Our goal is to maximize the utility function $f$ subject to our budget constraint, write down the three equations one derives from using the method of Lagrange multipliers.

(e) Solve the optimization problem. State both the optimal mix of nuts and the maximum utility.

(f) Suppose that your budget is allowed to change, so your solution to the above problem depends upon $M$; namely, you have three functions $(x(M), y(M), z(M))$. Recall that the *marginal utility of income* is the rate of change of your utility with respect to your budget. Show that the marginal utility of income is $\lambda$. Hint: This is always true, so you can either show this is true for the example above, or you can do this abstractly.

11-24. Find the global maximum and minimum to $f(x, y) = y^2 - 16x^2$ on the set $\{(x, y) : y \leq 1 - x^2, y \geq 0\}$.

11-25. The town of Newton-ville is on an elliptical island, which on the $xy$-plane can be described using the formula

$$x^2 + 2y^2 \leq 6.$$

A local topographer (a person who makes maps) observed that the altitude at any point $(x, y)$ in the town is given by the formula

$$A(x, y) = x^2 e^y.$$

Find the coordinates of the highest point(s) in town. Find the coordinates of the lowest point(s) in the town.

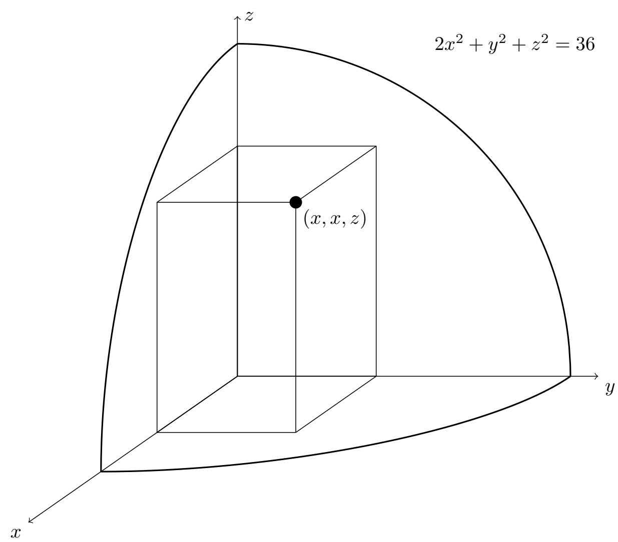
11-26. Consider the ellipsoid (a stretched sphere) with equation

$$2x^2 + y^2 + z^2 = 36,$$

and construct a box with a **square base** on the $xy$-plane, with one vertex at the origin $(0, 0, 0)$, and the diagonally opposite vertex on the surface of the ellipsoid (see figure).

Find the dimensions of the maximum-volume box.

[6] Note that we are assuming that $x, y, z$ are continuous, so this should be an equality rather than an inequality.




11-27. Recall that if $f$ is a function of one variable, its second order polynomial approximation at $a$ is

$$p_{2,a}(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2.$$

Now let $f(\mathbf{x})$ be a function of $n$-variables. The second order polynomial approximation to $f$ at $\mathbf{a} \in \mathbb{R}^n$ is

$$p_{2,a}(\mathbf{x}) = f(\mathbf{a}) + \nabla f(\mathbf{a})(\mathbf{x} - \mathbf{a}) + \frac{1}{2}(\mathbf{x} - \mathbf{a})^T H_f(\mathbf{a})(\mathbf{x} - \mathbf{a}).$$

Find the second order polynomial approximation to the function $f(x,y) = xe^y$ at $(a_1, a_2) = (2,3)^T$.

11-28. Find the extreme values of the function $f(x,y) = e^{xy}$ on the ellipse $2x^2 + 3y^2 \leq 6$.

11-29. Determine the integral of the function $f$ on each given rectangle $R$.

(a) $f(x,y) = x^2 + y$ on $R = [1,2] \times [2,3]$

(b) $f(x,y) = \sqrt{1+x+y}$ on $R = [0,9] \times [0,4]$

(c) $f(x,y) = x^5 y^3$ on $R = [0,2] \times [-2,2]$

(d) $f(x,y) = \frac{1}{x+y}$ on $R = [1,2] \times [1,2]$

11-30. Find the integral of $f$ over each area $\Omega$.

(a) $f(x,y) = 6y^2$ where $\Omega$ is the triangle whose vertices are $(-2,2), (0,0), (2,2)$.

(b) $f(x,y) = e^{x^2}$ where $\Omega$ is the region in the first quadrant bounded by $x = 4y, x = 0, x = 8, y = 0$.

(c) $f(x,y) = x + 2y$ where $\Omega$ is the region bounded by $y = 2x^2$ and $y = 1 + x^2$

(d) $f(x,y) = 1 + x + y$, $S = \{0 \leq x \leq 1, 0 \leq y \leq e^x\}$,

(e) $f(x,y) = (x-y)^2$, $S$ is the region bounded between $x^2$ and $x^3$,

(f) $f(x,y) = x^2y^2$, $S = \left\{-y^2 \leq x \leq y^2, 0 \leq y \leq 1\right\}$,

(g) $f(x,y) = xy$, $S$ the area bounded by the lines $y = x - 1$ and $y^2 = 2x + 6$, 11-31. Consider the integral

$$\int_0^4 \int_0^{\sqrt{4-y}} \frac{1}{\sqrt{12x - x^3}} \,\mathrm{d}x\,\mathrm{d}y.$$

(a) Sketch the region of integration.

(b) Rewrite the integral so it is of the form $\displaystyle\int_a^b \int_{f(x)}^{g(x)} \frac{1}{\sqrt{12x - x^3}} \,\mathrm{d}y\,\mathrm{d}x$ for appropriate choices of $a, b \in \mathbb{R}$ and functions $f, g$.

(c) Evaluate the integral you found in part (b).

11-32. Let $R$ be the region in $\mathbb{R}^2$ bounded by $y = 4 - x^2$ and $y = x + 2$. Let $f(x, y)$ be a function of two variables.

(a) Set up the integral $\displaystyle\iint_R f(x, y)\,\mathrm{d}A$ by integrating as $\mathrm{d}x\,\mathrm{d}y$. *Note:* If you accidentally reverse this, it's fine, since you'll do the opposite in part (b).

(b) Set up the integral $\displaystyle\iint_R f(x, y)\,\mathrm{d}A$ by integrating as $\mathrm{d}y\,\mathrm{d}x$.

(c) Let $f(x, y) = y$. Evaluate $\displaystyle\iint_R y\,\mathrm{d}A$ using either of the above integrals. Clearly indicate which method you are using.

11-33. Evaluate the integral of the following functions on the specifieddomain:

(a) $f(x, y, z) = y$ over the region bounded by the planes $x = 0$, $y = 0$, $z = 0$, and $2x+2y+z = 4$,

(b) $f(x, y, z) = z$ over the region bounded by $y^2 + z^2 = 9$, $x = 0$, $y = 3x$ and $z = 0$ in the first octant.

(c) $f(x, y, z) = 1$ over the region bounded by $y = x^2$, $z = 0$ and $y + z = 1$.

11-34. Determine the integral of $f$ on the given region $R$:

(a) $f(x, y) = xy$ on the region $R$ bounded by $y - x = 0, y - x = -1, x + 2y = 0, x + 2y = 6$

(b) $f(x, y) = x + y$ where $R$ is the trapezoid given by $(0, 0), (5, 0), (5/2, -5/2), (5/2, 5/2)$. *Hint:* Use the transformation $(u, v) = (y - x, y + x)$.

(c) $f(x, y) = \frac{1}{x+y}$ where $R$ is the region bounded by $x + y = 1$ and $x + y = 4$ in the first quadrant. *Hint:* Use the transformation $(x, y) = (u - uv, uv)$.

(d) $f(x, y) = x^2 y$ on the region $R$ bounded by $x + y^2 = 1$ and $x + 5y^2 = 5$.