Legal stuff

- Copyright © 2017 Joel Feldman and Andrew Rechnitzer

- In the near future this will be licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. You can view a copy of the license at http://creativecommons.org/licenses/by-nc-sa/4.0/.
Calculus is built on two operations — differentiation and integration.

- Differentiation — as we saw last term, differentiation allows us to compute and study the instantaneous rate of change of quantities. At its most basic it allows us to compute tangent lines and velocities, but it also led us to quite sophisticated applications including approximation of functions through Taylor polynomials and optimisation of quantities by studying critical and singular points.

- Integration — at its most basic, allows us to analyse the area under a curve. Of course, its application and importance extend far beyond areas and it plays a central role in solving differential equations.

It is not immediately obvious that these two topics are related to each other. However, as we shall see, they are indeed intimately linked.

**1.1 Definition of the Integral**

Arguably the easiest way to introduce integration is by considering the area between the graph of a given function and the $x$-axis, between two specific vertical lines — such as is shown in the figure above. We’ll follow this route by starting with a motivating example.
**A Motivating Example**

Let us find the area under the curve $y = e^x$ (and above the $x$-axis) for $0 \leq x \leq 1$. That is, the area of $\{(x, y) \mid 0 \leq y \leq e^x, 0 \leq x \leq 1\}$.

![Graph of $y = e^x$]

This area is equal to the “definite integral”

$$\text{Area} = \int_0^1 e^x \, dx$$

Do not worry about this notation or terminology just yet. We discuss it at length below. In different applications this quantity will have different interpretations — not just area. For example, if $x$ is time and $e^x$ is your velocity at time $x$, then we’ll see later (in Example 1.1.18) that the specified area is the net distance travelled between time 0 and time 1. After we finish with the example, we’ll mimic it to give a general definition of the integral $\int_a^b f(x) \, dx$.

**Example 1.1.1**

We wish to compute the area of $\{(x, y) \mid 0 \leq y \leq e^x, 0 \leq x \leq 1\}$. We know, from our experience with $e^x$ in differential calculus, that the curve $y = e^x$ is not easily written in terms of other simpler functions, so it is very unlikely that we would be able to write the area as a combination of simpler geometric objects such as triangles, rectangles or circles.

So rather than trying to write down the area exactly, our strategy is to approximate the area and then make our approximation more and more precise. We choose to approximate the area as a union of a large number of tall thin (vertical) rectangles. As we take more and more rectangles we get better and better approximations. Taking the limit as the number of rectangles goes to infinity gives the exact area.

As a warm up exercise, we’ll now just use four rectangles. In Example 1.1.2, below, we’ll consider an arbitrary number of rectangles and then take the limit as the number of rectangles goes to infinity. So

---

1. This should remind the reader of the approach taken to compute the slope of a tangent line way way back at the start of differential calculus.
2. Approximating the area in this way leads to a definition of integration that is called Riemann integration. This is the most commonly used approach to integration. However we could also approximate the area by using long thin horizontal strips. This leads to a definition of integration that is called Lebesgue integration. We will not be covering Lebesgue integration in these notes.
3. If we want to be more careful here, we should construct two approximations, one that is always a little smaller than the desired area and one that is a little larger. We can then take a limit using the Squeeze Theorem and arrive at the exact area. More on this later.
• subdivide the interval $0 \leq x \leq 1$ into 4 equal subintervals each of width $\frac{1}{4}$, and

• subdivide the area of interest into four corresponding vertical strips, as in the figure below.

The area we want is exactly the sum of the areas of all four strips.

Each of these strips is almost, but not quite, a rectangle. While the bottom and sides are fine (the sides are at right-angles to the base), the top of the strip is not horizontal. This is where we must start to approximate. We can replace each strip by a rectangle by just levelling off the top. But now we have to make a choice — at what height do we level off the top?

Consider, for example, the leftmost strip. On this strip, $x$ runs from 0 to $\frac{1}{4}$. As $x$ runs from 0 to $\frac{1}{4}$, the height $y$ runs from $e^0$ to $e^{1/4}$. It would be reasonable to choose the height of the approximating rectangle to be somewhere between $e^0$ and $e^{1/4}$. Which height should we choose? Well, actually it doesn’t matter. When we eventually take the limit of infinitely many approximating rectangles all of those different choices give exactly the same final answer. We’ll say more about this later.

In this example we’ll do two sample computations.

• For the first computation we approximate each slice by a rectangle whose height is the height of the left hand side of the slice.

  - On the first slice, $x$ runs from 0 to $\frac{1}{4}$, and the height $y$ runs from $e^0$, on the left hand side, to $e^{1/4}$, on the right hand side.
- So we approximate the first slice by the rectangle of height $e^0$ and width $1/4$, and hence of area $\frac{1}{4} e^0 = \frac{1}{4}$.
- On the second slice, $x$ runs from $1/4$ to $1/2$, and the height $y$ runs from $e^{1/4}$ and $e^{1/2}$.
- So we approximate the second slice by the rectangle of height $e^{1/4}$ and width $1/4$, and hence of area $\frac{1}{4} e^{1/4}$.
- And so on.
- All together, we approximate the area of interest by the sum of the areas of the four approximating rectangles, which is

$$[1 + e^{1/4} + e^{1/2} + e^{3/4}] \frac{1}{4} = 1.5124$$

- This particular approximation is called the “left Riemann sum approximation to $\int_0^1 e^x \, dx$ with 4 subintervals”. We’ll explain this terminology later.
- This particular approximation represents the shaded area in the figure on the left below. Note that, because $e^x$ increases as $x$ increases, this approximation is definitely smaller than the true area.

For the second computation we approximate each slice by a rectangle whose height is the height of the right hand side of the slice.

- On the first slice, $x$ runs from 0 to $1/4$, and the height $y$ runs from $e^0$, on the left hand side, to $e^{1/4}$, on the right hand side.
- So we approximate the first slice by the rectangle of height $e^{1/4}$ and width $1/4$, and hence of area $\frac{1}{4} e^{1/4}$.
- On the second slice, $x$ runs from $1/4$ to $1/2$, and the height $y$ runs from $e^{1/4}$ and $e^{1/2}$. 
- So we approximate the second slice by the rectangle of height \(e^{1/2}\) and width \(1/4\), and hence of area \(\frac{1}{4} e^{1/2}\).
- And so on.
- All together, we approximate the area of interest by the sum of the areas of the four approximating rectangles, which is

\[
\left[ e^{1/4} + e^{1/2} + e^{3/4} + e^1 \right] \frac{1}{4} = 1.9420
\]

- This particular approximation is called the “right Riemann sum approximation to \(\int_0^1 e^x \, dx\) with 4 subintervals”.
- This particular approximation represents the shaded area in the figure on the right above. Note that, because \(e^x\) increases as \(x\) increases, this approximation is definitely larger than the true area.

Now for the full computation that gives the exact area.

Recall that we wish to compute the area of \(\{(x, y) \mid 0 \leq y \leq e^x, 0 \leq x \leq 1\}\) and that our strategy is to approximate this area by the area of a union of a large number of very thin rectangles, and then take the limit as the number of rectangles goes to infinity. In Example 1.1.1, we used just four rectangles. Now we’ll consider a general number of rectangles, that we’ll call \(n\). Then we’ll take the limit \(n \to \infty\). So

- pick a natural number \(n\) and
- subdivide the interval \(0 \leq x \leq 1\) into \(n\) equal subintervals each of width \(1/n\), and
- subdivide the area of interest into corresponding thin strips, as in the figure below.

The area we want is exactly the sum of the areas of all of the thin strips.
Each of these strips is almost, but not quite, a rectangle. As in Example 1.1.1, the only problem is that the top is not horizontal. So we approximate each strip by a rectangle, just by levelling off the top. Again, we have to make a choice — at what height do we level off the top?

Consider, for example, the leftmost strip. On this strip, \(x\) runs from 0 to \(\frac{1}{n}\). As \(x\) runs from 0 to \(\frac{1}{n}\), the height \(y\) runs from \(e^0\) to \(e^{1/n}\). It would be reasonable to choose the height of the approximating rectangle to be somewhere between \(e^0\) and \(e^{1/n}\). Which height should we choose?

Well, as we said in Example 1.1.1, it doesn’t matter. We shall shortly take the limit \(n \to \infty\) and, in that limit, all of those different choices give exactly the same final answer. We won’t justify that statement in this example, but there will be an (optional) section shortly that provides the justification. For this example we just, arbitrarily, choose the height of each rectangle to be the height of the graph \(y = e^x\) at the smallest value of \(x\) in the corresponding strip. The figure on the left below shows the approximating rectangles when \(n = 4\) and the figure on the right shows the approximating rectangles when \(n = 8\).

Now we compute the approximating area when there are \(n\) strips.

- We approximate the leftmost strip by a rectangle of height \(e^0\). All of the rectangles have width \(\frac{1}{n}\). So the leftmost rectangle has area \(\frac{1}{n}e^0\).

- On strip number 2, \(x\) runs from \(\frac{1}{n}\) to \(\frac{2}{n}\). So the smallest value of \(x\) on strip number 2 is \(\frac{1}{n}\), and we approximate strip number 2 by a rectangle of height \(e^{1/n}\) and hence of area \(\frac{1}{n}e^{1/n}\).

- And so on.

- On the last strip, \(x\) runs from \(\frac{n-1}{n}\) to \(\frac{n}{n} = 1\). So the smallest value of \(x\) on the last strip is \(\frac{n-1}{n}\), and we approximate the last strip by a rectangle of height \(e^{(n-1)/n}\) and hence of area \(\frac{1}{n}e^{(n-1)/n}\).

\[\text{Notice that since } e^x \text{ is an increasing function, this choice of heights means that each of our rectangles is smaller than the strip it came from.}\]
The total area of all of the approximating rectangles is

\[
\text{Total approximating area} = \frac{1}{n} e^0 + \frac{1}{n} e^{1/n} + \frac{1}{n} e^{2/n} + \frac{1}{n} e^{3/n} + \cdots + \frac{1}{n} e^{(n-1)/n}
\]

\[
= \frac{1}{n} \left( 1 + e^{1/n} + e^{2/n} + e^{3/n} + \cdots + e^{(n-1)/n} \right)
\]

Now the sum in the brackets might look a little intimidating because of all the exponentials, but it actually has a pretty simple structure that can be easily seen if we rename \( e^{1/n} = r \). Then

- the first term is 1 = \( r^0 \) and
- the second term is \( e^{1/n} = r^1 \) and
- the third term is \( e^{2/n} = r^2 \) and
- the fourth term is \( e^{3/n} = r^3 \) and
- and so on and
- the last term is \( e^{(n-1)/n} = r^{n-1} \).

So

\[
\text{Total approximating area} = \frac{1}{n} \left( 1 + r + r^2 + \cdots + r^{n-1} \right)
\]

The sum in brackets is known as a geometric sum and satisfies a nice simple formula:

\[
1 + r + r^2 + \cdots + r^{n-1} = \frac{r^n - 1}{r - 1} \quad \text{provided } r \neq 1
\]

The derivation of the above formula is not too difficult. So let’s derive it in a little aside.

**Geometric sum**

Denote the sum as

\[
S = 1 + r + r^2 + \cdots + r^{n-1}
\]

Notice that if we multiply the whole sum by \( r \) we get back almost the same thing:

\[
rS = r \left( 1 + r + r^2 + \cdots + r^{n-1} \right) = r + r^2 + r^3 + \cdots + r^n
\]

This right hand side differs from the original sum \( S \) only in that

- the right hand side is missing the “1+” that \( S \) starts with and
• the right hand side has an extra "+r^n" at the end that does not appear in S.

That is

\[ rS = S - 1 + r^n \]

Moving this around a little gives

\[ (r - 1)S = (r^n - 1) \]

\[ S = \frac{r^n - 1}{r - 1} \]

as required. Notice that the last step in the manipulations only works providing \( r \neq 1 \) (otherwise we are dividing by zero).

>>> Back to approximating areas

Now we can go back to our area approximation armed with the above result about geometric sums.

Total approximating area = \[ \frac{1}{n} \left( 1 + r + r^2 + \ldots + r^{n-1} \right) \]

= \[ \frac{1}{n} \frac{r^n - 1}{r - 1} \]

= \[ \frac{1}{n} \frac{e^{n/n} - 1}{e^{1/n} - 1} \]

= \[ \frac{1}{n} \frac{e - 1}{e^{1/n} - 1} \]

To get the exact area\(^5\) all we need to do is make the approximation better and better by taking the limit \( n \to \infty \). The limit will look more familiar if we rename \( 1/n \) to \( X \). As \( n \) tends to infinity, \( X \) tends to 0, so

\[ \text{Area} = \lim_{n \to \infty} \frac{1}{n} \frac{e - 1}{e^{1/n} - 1} \]

= \( (e - 1) \lim_{n \to \infty} \frac{1/n}{e^{1/n} - 1} \)

= \( (e - 1) \lim_{X \to 0} \frac{X}{e^X - 1} \)

(with \( X = 1/n \))

Examining this limit we see that both numerator and denominator tend to zero as \( X \to 0 \), and so we cannot evaluate this limit by computing the limits of the numerator and denominator separately and then dividing the results. Despite this, the limit is not too hard to evaluate; here we give two ways:

\[ \begin{align*}
\text{Area} &= \lim_{n \to \infty} \frac{1}{n} \frac{e - 1}{e^{1/n} - 1} \\
&= (e - 1) \lim_{n \to \infty} \frac{1/n}{e^{1/n} - 1} \\
&= (e - 1) \lim_{X \to 0} \frac{X}{e^X - 1} \\
\end{align*} \]

\(^5\) We haven’t proved that this will give us the exact area, but it should be clear that taking this limit will give us a lower bound on the area. To complete things rigorously we also need an upper bound and the squeeze theorem. We do this in the next optional subsection.
• Perhaps the easiest way to compute the limit is by using l'Hôpital's rule\(^6\). Since both numerator and denominator go to zero, this is a \(0/0\) indeterminate form. Thus

\[
\lim_{x \to 0} \frac{X}{e^X - 1} = \lim_{x \to 0} \frac{\frac{d}{dx}X}{\frac{d}{dx}(e^X - 1)} = \lim_{x \to 0} \frac{1}{e^X} = 1
\]

• Another way\(^7\) to evaluate the same limit is to observe that it can be massaged into the form of the limit definition of the derivative. First notice that

\[
\lim_{x \to 0} \frac{X}{e^X - 1} = \left[ \lim_{x \to 0} \frac{e^X - 1}{X} \right]^{-1}
\]

provided this second limit exists and is nonzero. This second limit should look a little familiar:

\[
\lim_{x \to 0} \frac{e^X - 1}{X} = \lim_{x \to 0} \frac{e^X - e^0}{X - 0}
\]

which is just the definition of the derivative of \(e^x\) at \(x = 0\). Hence we have

\[
\lim_{x \to 0} \frac{X}{e^X - 1} = \left[ \lim_{x \to 0} \frac{e^X - e^0}{X - 0} \right]^{-1} = \left[ \frac{d}{dx}e^x \bigg|_{x=0} \right]^{-1} = [e^x]^{-1} = 1
\]

So, after this short aside into limits, we may now conclude that

\[
\text{Area} = (e - 1) \lim_{x \to 0} \frac{X}{e^X - 1} = e - 1
\]

1.1.1 ▶️ Optional — A more rigorous area computation

In Example 1.1.1 above we considered the area of the region \(\{(x, y) \mid 0 \leq y \leq e^x, 0 \leq x \leq 1\}\). We approximated that area by the area of a union of \(n\) thin rectangles. We then claimed that upon taking the number of rectangles to infinity, the approximation of the

---

\(^6\) If you do not recall L'Hôpital's rule and indeterminate forms then we recommend you skim over your differential calculus notes on the topic.

\(^7\) Say if you don’t recall l'Hôpital’s rule and have not had time to revise it.
area became the exact area. However we did not justify the claim. The purpose of this optional section is to make that calculation rigorous.

The broad set-up is the same. We divide the region up into \( n \) vertical strips, each of width \( \frac{1}{n} \) and we then approximate those strips by rectangles. However rather than an uncontrolled approximation, we construct two sets of rectangles — one set always smaller than the original area and one always larger. This then gives us lower and upper bounds on the area of the region. Finally we make use of the squeeze theorem\(^8\) to establish the result.

- To find our upper and lower bounds we make use of the fact that \( e^x \) is an increasing function. We know this because the derivative \( \frac{d}{dx} e^x = e^x \) is always positive. Consequently, the smallest and largest values of \( e^x \) on the interval \( a \leq x \leq b \) are \( e^a \) and \( e^b \), respectively.

- In particular, for \( 0 \leq x \leq \frac{1}{n} \), \( e^x \) takes values only between \( e^0 \) and \( e^{1/n} \). As a result, the first strip

\[
\left\{ (x, y) \mid 0 \leq x \leq \frac{1}{n}, \ 0 \leq y \leq e^x \right\}
\]

- contains the rectangle of \( 0 \leq x \leq \frac{1}{n}, \ 0 \leq y \leq e^0 \) (the lighter rectangle in the figure on the left below) and

- is contained in the rectangle \( 0 \leq x \leq \frac{1}{n}, \ 0 \leq y \leq e^{1/n} \) (the largest rectangle in the figure on the left below).

Hence

\[
\frac{1}{n} e^0 \leq \text{Area} \left\{ (x, y) \mid 0 \leq x \leq \frac{1}{n}, \ 0 \leq y \leq e^x \right\} \leq \frac{1}{n} e^{1/n}
\]

---

\(^8\) Recall that if we have 3 functions \( f(x), g(x), h(x) \) that satisfy \( f(x) \leq g(x) \leq h(x) \) and we know that \( \lim_{x \to a} f(x) = \lim_{x \to a} h(x) = L \) exists and is finite, then the squeeze theorem tells us that \( \lim_{x \to a} g(x) = L \).
• Similarly, for the second, third, …, last strips, as in the figure on the right above,

\[
\frac{1}{n} \leq \frac{1}{n} e^{1/n} \leq \text{Area}\{ (x, y) \mid 1/n \leq x \leq 2/n, 0 \leq y \leq e^x \} \leq \frac{1}{n} e^{2/n} \\
\frac{1}{n} \leq \frac{1}{n} e^{2/n} \leq \text{Area}\{ (x, y) \mid 2/n \leq x \leq 3/n, 0 \leq y \leq e^x \} \leq \frac{1}{n} e^{3/n} \\
\vdots \quad \vdots \\
\frac{1}{n} e^{(n-1)/n} \leq \text{Area}\{ (x, y) \mid (n-1)/n \leq x \leq n/n, 0 \leq y \leq e^x \} \leq \frac{1}{n} e^{n/n}
\]

• Adding these \( n \) inequalities together gives

\[
\frac{1}{n} \left( 1 + e^{1/n} + \ldots + e^{(n-1)/n} \right) \\
\quad \leq \text{Area}\{ (x, y) \mid 0 \leq x \leq 1, 0 \leq y \leq e^x \} \\
\quad \leq \frac{1}{n} \left( e^{1/n} + e^{2/n} + \ldots + e^{n/n} \right)
\]

• We can then recycle equation (1.1.3) with \( r = e^{1/n} \), so that \( r^n = \left( e^{1/n} \right)^n = e \). Thus we have

\[
\frac{1}{n} \frac{e - 1}{e^{1/n} - 1} \leq \text{Area}\{ (x, y) \mid 0 \leq x \leq 1, 0 \leq y \leq e^x \} \leq \frac{1}{n} \frac{e - 1}{e^{1/n} - 1}
\]

where we have used the fact that the upper bound is a simple multiple of the lower bound:

\[
\left( e^{1/n} + e^{2/n} + \ldots + e^{n/n} \right) = e^{1/n} \left( 1 + e^{1/n} + \ldots + e^{(n-1)/n} \right)
\]

• We now apply the squeeze theorem to the above inequalities. In particular, the limits of the lower and upper bounds are

\[
\lim_{n \to \infty} \frac{1}{n} \frac{e - 1}{e^{1/n} - 1} = (e - 1) \quad \lim_{X \to 0} \frac{X}{e^X - 1} = e - 1
\]

(by l’Hôpital’s rule) and

\[
\lim_{n \to \infty} \frac{1}{n} \frac{e - 1}{e^{1/n} - 1} = (e - 1) \lim_{X \to 0} \frac{X e^X}{e^X - 1} \\
= (e - 1) \lim_{X \to 0} e^X \cdot \lim_{X \to 0} \frac{X}{e^X - 1} \\
= (e - 1) \cdot 1 \cdot 1
\]

Thus, since the exact area is trapped between the lower and upper bounds, the squeeze theorem then implies that

\[
\text{Exact area} = e - 1.
\]
1.1.2 ▶ Summation notation

As you can see from the above example (and the more careful rigorous computation), our discussion of integration will involve a fair bit of work with sums of quantities. To this end, we make a quick aside into summation notation. While one can work through the material below without this notation, proper summation notation is well worth learning, so we advise the reader to persevere.

Writing out the summands explicitly can become quite impractical — for example, say we need the sum of the first 11 squares:

$$1 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2 + 7^2 + 8^2 + 9^2 + 10^2 + 11^2$$

This becomes tedious. Where the pattern is clear, we will often skip the middle few terms and instead write

$$1 + 2^2 + \cdots + 11^2.$$ 

A far more precise way to write this is using \(\Sigma\) (capital-sigma) notation. For example, we can write the above sum as

$$\sum_{k=1}^{11} k^2$$

This is read as

The sum from \(k\) equals 1 to 11 of \(k^2\).

More generally

**Notation 1.1.4.**

Let \(m \leq n\) be integers and let \(f(x)\) be a function defined on the integers. Then we write

$$\sum_{k=m}^{n} f(k)$$

to mean the sum of \(f(k)\) for \(k\) from \(m\) to \(n\):

$$f(m) + f(m+1) + f(m+2) + \cdots + f(n-1) + f(n).$$

Similarly we write

$$\sum_{i=m}^{n} a_i$$

to mean

$$a_m + a_{m+1} + a_{m+2} + \cdots + a_{n-1} + a_n$$

for some set of coefficients \(\{a_m, \ldots, a_n\}\).
Consider the example

$$\sum_{k=3}^{7} \frac{1}{k^2} = \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \frac{1}{6^2} + \frac{1}{7^2}$$

It is important to note that the right hand side of this expression evaluates to a number; it does not contain “$k$”. The summation index $k$ is just a “dummy” variable and it does not have to be called $k$. For example

$$\sum_{k=3}^{7} \frac{1}{k^2} = \sum_{i=3}^{7} \frac{1}{i^2} = \sum_{j=3}^{7} \frac{1}{j^2} = \sum_{\ell=3}^{7} \frac{1}{\ell^2}$$

Also the summation index has no meaning outside the sum. For example

$$k \sum_{k=3}^{7} \frac{1}{k^2}$$

has no mathematical meaning; it is gibberish.

A sum can be represented using summation notation in many different ways. If you are unsure as to whether or not two summation notations represent the same sum, just write out the first few terms and the last couple of terms. For example,

$$\sum_{m=3}^{15} \frac{1}{m^2} = \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \cdots + \frac{1}{14^2} + \frac{1}{15^2}$$

$$\sum_{m=4}^{16} \frac{1}{(m-1)^2} = \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \cdots + \frac{1}{14^2} + \frac{1}{15^2}$$

are equal.

Here is a theorem that gives a few rules for manipulating summation notation.

**Theorem 1.1.5 (Arithmetic of Summation Notation).**

Let $n \geq m$ be integers. Then for all real numbers $c$ and $a_i, b_i, m \leq i \leq n$.

(a) $\sum_{i=m}^{n} ca_i = c \left( \sum_{i=m}^{n} a_i \right)$

(b) $\sum_{i=m}^{n} (a_i + b_i) = \left( \sum_{i=m}^{n} a_i \right) + \left( \sum_{i=m}^{n} b_i \right)$

(c) $\sum_{i=m}^{n} (a_i - b_i) = \left( \sum_{i=m}^{n} a_i \right) - \left( \sum_{i=m}^{n} b_i \right)$

---

9 Some careful addition shows it is $\frac{46181}{176400}$. 
**Proof.** We can prove this theorem by just writing out both sides of each equation, and observing that they are equal, by the usual laws of arithmetic. For example, for the first equation, the left and right hand sides are

\[ \sum_{i=m}^{n} ca_i = ca_m + ca_{m+1} + \cdots + ca_n \quad \text{and} \quad c \left( \sum_{i=m}^{n} a_i \right) = c(a_m + a_{m+1} + \cdots + a_n) \]

They are equal by the usual distributive law. The “distributive law” is the fancy name for \( c(a + b) = ca + cb \).

Not many sums can be computed exactly. Here are some that can. The first few are used a lot.

---

**Theorem 1.1.6.**

(a) \( \sum_{i=0}^{n} ar^i = a \frac{1-r^{n+1}}{1-r} \), for all real numbers \( a \) and \( r \neq 1 \) and all integers \( n \geq 0 \).

(b) \( \sum_{i=1}^{n} 1 = n \), for all integers \( n \geq 1 \).

(c) \( \sum_{i=1}^{n} i = \frac{1}{2} n(n + 1) \), for all integers \( n \geq 1 \).

(d) \( \sum_{i=1}^{n} i^2 = \frac{1}{6} n(n + 1)(2n + 1) \), for all integers \( n \geq 1 \).

(e) \( \sum_{i=1}^{n} i^3 = \left[ \frac{1}{2} \frac{1}{2} n(n + 1) \right]^2 \), for all integers \( n \geq 1 \).

---

10 Since all the sums are finite, this isn’t too hard. More care must be taken when the sums involve an infinite number of terms. We will examine this in Chapter 3.

11 Of course, any finite sum can be computed exactly — just sum together the terms. What we mean by “computed exactly” in this context, is that we can rewrite the sum as a simple, and easily evaluated, formula involving the terminals of the sum. For example

\[ \sum_{k=m}^{n} r^k = \frac{r^{n+1} - r^m}{r - 1} \quad \text{provided} \quad r \neq 1 \]

No matter what finite integers we choose for \( m \) and \( n \), we can quickly compute the sum in just a few arithmetic operations. On the other hand, the sums,

\[ \sum_{k=m}^{n} \frac{1}{k}, \quad \sum_{k=m}^{n} \frac{1}{k^2} \]

cannot be expressed in such clean formulas (though you can rewrite them quite cleanly using integrals). To explain more clearly we would need to go into a more detailed and careful discussion that is beyond the scope of this course.
Proof of Theorem 1.1.6 (Optional)

Proof. (a) The first sum is
\[ \sum_{i=0}^{n} ar^i = ar^0 + ar^1 + ar^2 + \cdots + ar^n \]
which is just the left hand side of equation (1.1.3), with \( n \) replaced by \( n + 1 \) and then multiplied by \( a \).

(b) The second sum is just \( n \) copies of 1 added together, so of course the sum is \( n \).

(c) The third and fourth sums are discussed in the appendix of the CLP notes for mathematics 100 and 180. In that discussion certain “tricks” are used to compute the sums with only simple arithmetic. Those tricks do not easily generalise to the fifth sum.

(c') Instead of repeating that appendix, we’ll derive the third sum using a trick that generalises to the fourth and fifth sums (and also to higher powers). The trick uses the generating function \[ S(x) = 1 + x + x^2 + \cdots + x^n = \frac{x^{n+1} - 1}{x - 1} \] Notice that this is just the geometric sum given by equation 1.1.3 with \( n \) replaced by \( n + 1 \).

Now, consider the limit
\[ \lim_{x \to 1} S(x) = \lim_{x \to 1} \left( 1 + x + x^2 + \cdots + x^n \right) = n + 1 \]
but also
\[ = \lim_{x \to 1} \frac{x^{n+1} - 1}{x - 1} \]
now use l’Hôpital’s rule
\[ = \lim_{x \to 1} \frac{(n + 1)x^n}{1} = n + 1. \]

This is not so hard (or useful). But now consider the derivative of \( S(x) \):
\[ S'(x) = 1 + 2x + 3x^2 + \cdots + nx^{n-1} \]
\[ = \frac{d}{dx} \left[ \frac{x^{n+1} - 1}{x - 1} \right] \quad \text{use the quotient rule} \]
\[ = \frac{(x - 1) \cdot (n + 1)x^n - (x^{n+1} - 1) \cdot 1}{(x - 1)^2} \quad \text{now clean it up} \]
\[ = \frac{nx^{n+1} - (n + 1)x^n + 1}{(x - 1)^2}. \]

---

12 Generating functions are frequently used in mathematics to analyse sequences and series, but are beyond the scope of the course. The interested reader should take a look at “Generatingfunctionology” by Herb Wilf. It is an excellent book and is also free to download.
Here if we take the limit of the above expression as $x \to 1$ we recover

\[
\lim_{x \to 1} S'(x) = 1 + 2 + 3 + \cdots + n
\]

\[
= \lim_{x \to 1} \frac{nx^{n+1} - (n + 1)x^n + 1}{(x - 1)^2}
\]

now use l’Hôpital’s rule

\[
= \lim_{x \to 1} \frac{n(n + 1)x^n - n(n + 1)x^{n-1}}{2(x - 1)}
\]
l’Hôpital’s rule again

\[
= \lim_{x \to 1} \frac{n^2(n + 1)x^{n-1} - n(n + 1)(n - 1)x^{n-2}}{2}
\]

\[
= \frac{n^2(n + 1) - n(n - 1)(n + 1)}{2} = \frac{n(n + 1)}{2}
\]

as required. This computation can be done without l’Hôpital’s rule, but the manipulations required are a fair bit messier.

(d) The derivation of the fourth and fifth sums is similar to, but even more tedious than, that of the third sum. One takes two or three derivatives of the generating functional.

\[
\square
\]

1.1.3 The Definition of the Definite Integral

In this section we give a definition of the definite integral $\int_a^b f(x)\,dx$ generalising the machinery we used in Example 1.1.1. But first some terminology and a couple of remarks to better motivate the definition.

**Notation 1.1.8.**

The symbol $\int_a^b f(x)\,dx$ is read “the definite integral of the function $f(x)$ from $a$ to $b$”. The function $f(x)$ is called the integrand of $\int_a^b f(x)\,dx$ and $a$ and $b$ are called the limits of integration. The interval $a \leq x \leq b$ is called the interval of integration and is also called the domain of integration.

Before we explain more precisely what the definite integral actually is, a few remarks (actually — a few interpretations) are in order.

- If $f(x) \geq 0$ and $a \leq b$, one interpretation of the symbol $\int_a^b f(x)\,dx$ is “the area of the region $\{ (x,y) \mid a \leq x \leq b, \, 0 \leq y \leq f(x) \}$.”
In this way we can rewrite the area in Example 1.1.1 as the definite integral $\int_{0}^{1} e^x \, dx$.

• This interpretation breaks down when either $a > b$ or $f(x)$ is not always positive, but it can be repaired by considering “signed areas”.

• If $a \leq b$, but $f(x)$ is not always positive, one interpretation of $\int_{a}^{b} f(x) \, dx$ is “the signed area between $y = f(x)$ and the $x$–axis for $a \leq x \leq b$”. For “signed area” (which is also called the “net area”), areas above the $x$–axis count as positive while areas below the $x$–axis count as negative. In the example below, we have the graph of the function

$$f(x) = \begin{cases} 
-1 & \text{if } 1 \leq x \leq 2 \\
2 & \text{if } 2 < x \leq 4 \\
0 & \text{otherwise}
\end{cases}$$

The $2 \times 2$ shaded square above the $x$–axis has signed area $+2 \times 2 = +4$. The $1 \times 1$ shaded square below the $x$–axis has signed area $-1 \times 1 = -1$. So, for this $f(x)$,

$$\int_{0}^{5} f(x) \, dx = +4 - 1 = 3$$

• We’ll come back to the case $b < a$ later.
We’re now ready to define $\int_a^b f(x) \, dx$. The definition is a little involved, but essentially mimics what we did in Example 1.1.1 (which is why we did the example before the definition). The main differences are that we replace the function $e^x$ by a generic function $f(x)$ and we replace the interval from 0 to 1 by the generic interval from $a$ to $b$.

- We start by selecting any natural number $n$ and subdividing the interval from $a$ to $b$ into $n$ equal subintervals. Each subinterval has width $\frac{b-a}{n}$.

- Just as was the case in Example 1.1.1 we will eventually take the limit as $n \to \infty$, which squeezes the width of each subinterval down to zero.

- For each integer $0 \leq i \leq n$, define $x_i = a + i \cdot \frac{b-a}{n}$. Note that this means that $x_0 = a$ and $x_n = b$. It is worth keeping in mind that these numbers $x_i$ do depend on $n$ even though our choice of notation hides this dependence.

- Subinterval number $i$ is $x_{i-1} \leq x \leq x_i$. In particular, on the first subinterval, $x$ runs from $x_0 = a$ to $x_1 = a + \frac{b-a}{n}$. On the second subinterval, $x$ runs from $x_1$ to $x_2 = a + 2 \frac{b-a}{n}$.

- On each subinterval we now pick $x_{i,n}$ between $x_{i-1}$ and $x_i$. We then approximate $f(x)$ on the $i$th subinterval by the constant function $y = f(x_{i,n})$. We include $n$ in the subscript to remind ourselves that these numbers depend on $n$.

Geometrically, we’re approximating the region

$$\{ (x, y) \mid x \text{ is between } x_{i-1} \text{ and } x_i, \text{ and } y \text{ is between } 0 \text{ and } f(x) \}$$

by the rectangle

$$\{ (x, y) \mid x \text{ is between } x_{i-1} \text{ and } x_i, \text{ and } y \text{ is between } 0 \text{ and } f(x_{i,n}) \}$$

13 We’ll eventually allow $a$ and $b$ to be any two real numbers, not even requiring $a < b$. But it is easier to start off assuming $a < b$, and that’s what we’ll do.
In Example 1.1.1 we chose \( x_{i,n}^* = x_{i-1} \) and so we approximated the function \( e^x \) on each subinterval by the value it took at the leftmost point in that subinterval.

- So, when there are \( n \) subintervals our approximation to the signed area between the curve \( y = f(x) \) and the \( x \)-axis, with \( x \) running from \( a \) to \( b \), is

\[
\sum_{i=1}^{n} f(x_{i,n}^*) \cdot \frac{b-a}{n}
\]

We interpret this as the signed area since the summands \( f(x_{i,n}^*) \cdot \frac{b-a}{n} \) need not be positive.

- Finally we define the definite integral by taking the limit of this sum as \( n \to \infty \).

Oof! This is quite an involved process, but we can now write down the definition we need.

**Definition 1.1.9.**

Let \( a \) and \( b \) be two real numbers and let \( f(x) \) be a function that is defined for all \( x \) between \( a \) and \( b \). Then we define

\[
\int_{a}^{b} f(x) \, dx = \lim_{n \to \infty} \sum_{i=1}^{n} f(x_{i,n}^*) \cdot \frac{b-a}{n}
\]

when the limit exists and takes the same value for all choices of the \( x_{i,n}^* \)'s. In this case, we say that \( f \) is integrable on the interval from \( a \) to \( b \).

Of course, it is not immediately obvious when this limit should exist. Thankfully it is easier for a function to be “integrable” than it is for it to be “differentiable”.
Theorem 1.1.10.

Let $f(x)$ be a function on the interval $[a, b]$. If

- $f(x)$ is continuous on $[a, b]$, or
- $f(x)$ has a finite number of jump discontinuities on $[a, b]$ (and is otherwise continuous)

then $f(x)$ is integrable on $[a, b]$.

We will not justify this theorem. But a slightly weaker statement is proved in (the optional) Section 1.1.6. Of course this does not tell us how to actually evaluate any definite integrals — but we will get to that in time.

Some comments:

- Note that, in Definition 1.1.9, we allow $a$ and $b$ to be any two real numbers. We do not require that $a < b$. That is, even when $a > b$, the symbol $\int_a^b f(x) \, dx$ is still defined by the formula of Definition 1.1.9. We’ll get an interpretation for $\int_a^b f(x) \, dx$, when $a > b$, later.

- It is important to note that the definite integral $\int_a^b f(x) \, dx$ represents a number, not a function of $x$. The integration variable $x$ is another “dummy” variable, just like the summation index $i$ in $\sum_{i=m}^n a_i$ (see Section 1.1.2). The integration variable does not have to be called $x$. For example

$$\int_a^b f(x) \, dx = \int_a^b f(t) \, dt = \int_a^b f(u) \, du$$

Just as with summation variables, the integration variable $x$ has no meaning outside of $f(x) \, dx$. For example

$$x \int_0^1 e^x \, dx \quad \text{and} \quad \int_0^x e^x \, dx$$

are both gibberish.

The sum inside definition 1.1.9 is named after Bernhard Riemann\(^{14}\) who made the first rigorous definition of the definite integral and so placed integral calculus on rigorous footings.

\(^{14}\) Bernhard Riemann was a 19th century German mathematician who made extremely important contributions to many different areas of mathematics — far too many to list here. Arguably two of the most important (after Riemann sums) are now called Riemann surfaces and the Riemann hypothesis (he didn’t name them after himself).
1.1 Definition of the Integral

The sum inside definition 1.1.9

\[ \sum_{i=1}^{n} f(x_{i,n}^*) \frac{b-a}{n} \]

is called a Riemann sum. It is also often written as

\[ \sum_{i=1}^{n} f(x_i^*) \Delta x \]

where \( \Delta x = \frac{b-a}{n} \).

- If we choose each \( x_{i,n}^* = x_{i-1} = a + (i - 1) \frac{b-a}{n} \) to be the left hand end point of the \( i^{th} \) interval, \([x_{i-1}, x_i]\), we get the approximation

\[ \sum_{i=1}^{n} f \left( a + (i - 1) \frac{b-a}{n} \right) \frac{b-a}{n} \]

which is called the “left Riemann sum approximation to \( \int_{a}^{b} f(x) \, dx \) with \( n \) subintervals”. This is the approximation used in Example 1.1.1.

- In the same way, if we choose \( x_{i,n}^* = x_i = a + i \frac{b-a}{n} \) we obtain the approximation

\[ \sum_{i=1}^{n} f \left( a + i \frac{b-a}{n} \right) \frac{b-a}{n} \]

which is called the “right Riemann sum approximation to \( \int_{a}^{b} f(x) \, dx \) with \( n \) subintervals”. The word “right” signifies that, on each subinterval \([x_{i-1}, x_i]\) we approximate \( f \) by its value at the right-hand end–point, \( x_i = a + i \frac{b-a}{n} \), of the subinterval.

- A third commonly used approximation is

\[ \sum_{i=1}^{n} f \left( a + (i - 1/2) \frac{b-a}{n} \right) \frac{b-a}{n} \]

which is called the “midpoint Riemann sum approximation to \( \int_{a}^{b} f(x) \, dx \) with \( n \) subintervals”. The word “midpoint” signifies that, on each subinterval \([x_{i-1}, x_i]\) we approximate \( f \) by its value at the midpoint, \( \frac{x_{i-1} + x_i}{2} = a + (i - 1/2) \frac{b-a}{n} \), of the subinterval.

In order to compute a definite integral using Riemann sums we need to be able to
compute the limit of the sum as the number of summands goes to infinity. This approach is not always feasible and we will soon arrive at other means of computing definite integrals based on antiderivatives. However, Riemann sums also provide us with a good means of approximating definite integrals — if we take $n$ to be a large, but finite, integer, then the corresponding Riemann sum can be a good approximation of the definite integral. Under certain circumstances this can be strengthened to give rigorous bounds on the integral. Let us revisit Example 1.1.1.

Example 1.1.12

Let’s say we are again interested in the integral $\int_{0}^{1} e^x \, dx$. We can follow the same procedure as we used previously to construct Riemann sum approximations. However since the integrand $f(x) = e^x$ is an increasing function, we can make our approximations into upper and lower bounds without much extra work.

More precisely, we approximate $f(x)$ on each subinterval $x_{i-1} \leq x \leq x_i$

- by its smallest value on the subinterval, namely $f(x_{i-1})$, when we compute the left Riemann sum approximation and

- by its largest value on the subinterval, namely $f(x_i)$, when we compute the right Riemann sum approximation.

This is illustrated in the two figures below. The shaded region in the left hand figure is the left Riemann sum approximation and the shaded region in the right hand figure is the right Riemann sum approximation.

We can see that exactly because $f(x)$ is increasing, the left Riemann sum describes an area smaller than the definite integral while the right Riemann sum gives an area larger.\(^{15}\) than the integral.

When we approximate the integral $\int_{0}^{1} e^x \, dx$ using $n$ subintervals, then, on interval number $i$,

- $x$ runs from $\frac{i-1}{n}$ to $\frac{i}{n}$ and

15 When a function is decreasing the situation is reversed — the left Riemann sum is always larger than the integral while the right Riemann sum is smaller than the integral. For more general functions that both increase and decrease it is perhaps easiest to study each increasing (or decreasing) interval separately.
• \( y = e^x \) runs from \( e^{(i-1)/n} \), when \( x \) is at the left hand end point of the interval, to \( e^{i/n} \), when \( x \) is at the right hand end point of the interval.

Consequently, the left Riemann sum approximation to \( \int_0^1 e^x \, dx \) is \( \sum_{i=1}^{n} e^{(i-1)/n} \frac{1}{n} \) and the right Riemann sum approximation is \( \sum_{i=1}^{n} e^{i/n} \frac{1}{n} \). So

\[
\sum_{i=1}^{n} e^{(i-1)/n} \frac{1}{n} \leq \int_0^1 e^x \, dx \leq \sum_{i=1}^{n} e^{i/n} \frac{1}{n}
\]

Thus \( L_n = \sum_{i=1}^{n} e^{(i-1)/n} \frac{1}{n} \), which for any \( n \) can be evaluated by computer, is a lower bound on the exact value of \( \int_0^1 e^x \, dx \) and \( R_n = \sum_{i=1}^{n} e^{i/n} \frac{1}{n} \), which for any \( n \) can also be evaluated by computer, is an upper bound on the exact value of \( \int_0^1 e^x \, dx \). For example, when \( n = 1000 \), \( L_n = 1.7174 \) and \( R_n = 1.7191 \) (both to four decimal places) so that, again to four decimal places,

\[
1.7174 \leq \int_0^1 e^x \, dx \leq 1.7191
\]

Recall that the exact value is \( e - 1 = 1.718281828 \ldots \).

1.1.4 Using Known Areas to Evaluate Integrals

One of the main aims of this course is to build up general machinery for computing definite integrals (as well as interpreting and applying them). We shall start on this soon, but not quite yet. We have already seen one concrete, if laborious, method for computing definite integrals — taking limits of Riemann sums as we did in Example 1.1.1. A second method, which will work for some special integrands, works by interpreting the definite integral as “signed area”. This approach will work nicely when the area under the curve decomposes into simple geometric shapes like triangles, rectangles and circles. Here are some examples of this second method.

Example 1.1.13

The integral \( \int_a^b 1 \, dx \) (which is also written as just \( \int_a^b \, dx \)) is the area of the shaded rectangle (of width \( b - a \) and height 1) in the figure on the right below. So

\[
\int_a^b \, dx = (b - a) \times (1) = b - a
\]
Let \( b > 0 \). The integral \( \int_0^b x\,dx \) is the area of the shaded triangle (of base \( b \) and of height \( b \)) in the figure on the right below. So

\[
\int_0^b x\,dx = \frac{1}{2} b \times b = \frac{b^2}{2}
\]

The integral \( \int_{-b}^0 x\,dx \) is the signed area of the shaded triangle (again of base \( b \) and of height \( b \)) in the figure on the right below. So

\[
\int_{-b}^0 x\,dx = -\frac{b^2}{2}
\]

Notice that it is very easy to extend this example to the integral \( \int_0^b cx\,dx \) for any real numbers \( b, c > 0 \) and find

\[
\int_0^b cx\,dx = \frac{c}{2} b^2.
\]

In this example, we shall evaluate \( \int_{-1}^1 (1 - |x|)\,dx \). Recall that

\[
|x| = \begin{cases} 
-x & \text{if } x \leq 0 \\
-x & \text{if } x \geq 0
\end{cases}
\]

so that

\[
1 - |x| = \begin{cases} 
1 + x & \text{if } x \leq 0 \\
1 - x & \text{if } x \geq 0
\end{cases}
\]

To picture the geometric figure whose area the integral represents observe that

- at the left hand end of the domain of integration \( x = -1 \) and the integrand \( 1 - |x| = 1 - | -1 | = 1 - 1 = 0 \) and
• as \( x \) increases from \(-1\) towards 0, the integrand \( 1 - |x| = 1 + x \) increases linearly, until
• when \( x \) hits 0 the integrand hits \( 1 - |x| = 1 - |0| = 1 \) and then
• as \( x \) increases from 0, the integrand \( 1 - |x| = 1 - x \) decreases linearly, until
• when \( x \) hits \(+1\), the right hand end of the domain of integration, the integrand hits \( 1 - |x| = 1 - |1| = 0 \).

So the integral \( \int_{-1}^{1} (1 - |x|) \, dx \) is the area of the shaded triangle (of base 2 and of height 1) in the figure on the right below and

\[
\int_{-1}^{1} (1 - |x|) \, dx = \frac{1}{2} \times 2 \times 1 = 1
\]

Example 1.1.15

---

Example 1.1.16

The integral \( \int_{0}^{1} \sqrt{1 - x^2} \, dx \) has integrand \( f(x) = \sqrt{1 - x^2} \). So it represents the area under \( y = \sqrt{1 - x^2} \) with \( x \) running from 0 to 1. But we may rewrite

\[
y = \sqrt{1 - x^2} \quad \text{as} \quad x^2 + y^2 = 1, y \geq 0
\]

But this is the (implicit) equation for a circle — the extra condition that \( y \geq 0 \) makes it the equation for the semi-circle centred at the origin with radius 1 lying on and above the \( x \)-axis. Thus the integral represents the area of the quarter circle of radius 1, as shown in the figure on the right below. So

\[
\int_{0}^{1} \sqrt{1 - x^2} \, dx = \frac{1}{4} \pi (1)^2 = \frac{\pi}{4}
\]

Example 1.1.16

---

Example 1.1.17

This next one is a little trickier and relies on us knowing the symmetries of the sine function.

The integral \( \int_{-\pi}^{\pi} \sin x \, dx \) is the signed area of the shaded region in the figure on the right.
below. It naturally splits into two regions, one on either side of the y-axis. We don’t know the formula for the area of either of these regions (yet), however the two regions are very nearly the same. In fact, the part of the shaded region below the x-axis is exactly the reflection, in the x-axis, of the part of the shaded region above the x-axis. So the signed area of part of the shaded region below the x-axis is the negative of the signed area of part of the shaded region above the x-axis and

\[ \int_{-\pi}^{\pi} \sin x \, dx = 0 \]

1.1.5 Another Interpretation for Definite Integrals

So far, we have only a single interpretation\(^{16}\) for definite integrals — namely areas under graphs. In the following example, we develop a second interpretation.

Suppose that a particle is moving along the x-axis and suppose that at time \( t \) its velocity is \( v(t) \) (with \( v(t) > 0 \) indicating rightward motion and \( v(t) < 0 \) indicating leftward motion). What is the change in its x-coordinate between time \( a \) and time \( b > a \)?

We’ll work this out using a procedure similar to our definition of the integral. First pick a natural number \( n \) and divide the time interval from \( a \) to \( b \) into \( n \) equal subintervals, each of width \( \frac{b-a}{n} \). We are working our way towards a Riemann sum (as we have done several times above) and so we will eventually take the limit \( n \to \infty \).

- The first time interval runs from \( a \) to \( a + \frac{b-a}{n} \). If we think of \( n \) as some large number, the width of this interval, \( \frac{b-a}{n} \) is very small and over this time interval, the velocity does not change very much. Hence we can approximate the velocity over the first subinterval as being essentially constant at its value at the start of the time interval — \( v(a) \). Over the subinterval the x-coordinate changes by velocity times time, namely \( v(a) \cdot \frac{b-a}{n} \).

- Similarly, the second interval runs from time \( a + \frac{b-a}{n} \) to time \( a + 2\frac{b-a}{n} \). Again, we can assume that the velocity does not change very much and so we can approximate

\(^{16}\) If this were the only interpretation then integrals would be a nice mathematical curiosity and unlikely to be the core topic of a large first year mathematics course.
the velocity as being essentially constant at its value at the start of the subinterval — namely \( v(a + \frac{b-a}{n}) \). So during the second subinterval the particle’s \( x \)-coordinate changes by approximately \( v(a + \frac{b-a}{n}) \frac{b-a}{n} \).

- In general, time subinterval number \( i \) runs from \( a + (i-1) \frac{b-a}{n} \) to \( a + i \frac{b-a}{n} \) and during this subinterval the particle’s \( x \)-coordinate changes, essentially, by

\[
v(a + (i-1) \frac{b-a}{n}) \frac{b-a}{n} \]

So the net change in \( x \)-coordinate from time \( a \) to time \( b \) is approximately

\[
v(a) \frac{b-a}{n} + v(a + \frac{b-a}{n}) \frac{b-a}{n} + \cdots + v(a + (i-1) \frac{b-a}{n}) \frac{b-a}{n} + \cdots + v(a + (n-1) \frac{b-a}{n}) \frac{b-a}{n} = \sum_{i=1}^{n} v(a + (i-1) \frac{b-a}{n}) \frac{b-a}{n}
\]

This exactly the left Riemann sum approximation to the integral of \( v \) from \( a \) to \( b \) with \( n \) subintervals. The limit as \( n \to \infty \) is exactly the definite integral \( \int_{a}^{b} v(t) \, dt \). Following tradition, we have called the (dummy) integration variable \( t \) rather than \( x \) to remind us that it is time that is running from \( a \) to \( b \).

The conclusion of the above discussion is that if a particle is moving along the \( x \)-axis and its \( x \)-coordinate and velocity at time \( t \) are \( x(t) \) and \( v(t) \), respectively, then, for all \( b > a \),

\[
x(b) - x(a) = \int_{a}^{b} v(t) \, dt.
\]

Example 1.1.18

1.1.6 Optional — careful definition of the integral

In this optional section we give a more mathematically rigorous definition of the definite integral \( \int_{a}^{b} f(x) \, dx \). Some textbooks use a sneakier, but equivalent, definition. The integral will be defined as the limit of a family of approximations to the area between the graph of \( y = f(x) \) and the \( x \)-axis, with \( x \) running from \( a \) to \( b \). We will then show conditions under-which this limit is guaranteed to exist. We should state up front that these conditions are more restrictive than is strictly necessary — this is done so as to keep the proof accessible.

The family of approximations needed is slightly more general than that used to define Riemann sums in the previous sections, though it is quite similar. The main difference is that we do not require that all the subintervals have the same size.
• We start by selecting a positive integer \( n \). As was the case previously, this will be the number of subintervals used in the approximation and eventually we will take the limit as \( n \to \infty \).

• Now subdivide the interval from \( a \) to \( b \) into \( n \) subintervals by selecting \( n + 1 \) values of \( x \) that obey

\[
a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.
\]

The subinterval number \( i \) runs from \( x_{i-1} \) to \( x_i \). This formulation does not require the subintervals to have the same size. However we will eventually require that the widths of the subintervals shrink towards zero as \( n \to \infty \).

• Then for each subinterval we select a value of \( x \) in that interval. That is, for \( i = 1, 2, \ldots, n \), choose \( x_i^* \) satisfying \( x_{i-1} \leq x_i^* \leq x_i \). We will use these values of \( x \) to help approximate \( f(x) \) on each subinterval.

• The area between the graph of \( y = f(x) \) and the \( x \)-axis, with \( x \) running from \( x_{i-1} \) to \( x_i \), i.e. the contribution, \( \int_{x_{i-1}}^{x_i} f(x) \, dx \), from interval number \( i \) to the integral, is approximated by the area of a rectangle. The rectangle has width \( x_i - x_{i-1} \) and height \( f(x_i^*) \).

\[
f(x_i^*) \quad \text{rectangle}
\]
Thus the approximation to the integral, using all \( n \) subintervals, is
\[
\int_a^b f(x) \, dx \approx f(x_1^+) [x_1 - x_0] + f(x_2^+) [x_2 - x_1] + \cdots + f(x_n^+) [x_n - x_{n-1}]
\]

Of course every different choice of \( n \) and \( x_1, x_2, \ldots, x_n \) and \( x_1^+, x_2^+, \ldots, x_n^+ \) gives a different approximation. So to simplify the discussion that follows, let us denote a particular choice of all these numbers by \( P \):
\[
P = (n, x_1, x_2, \ldots, x_{n-1}, x_1^+, x_2^+, \ldots, x_n^+)
\]
Similarly let us denote the resulting approximation by \( \mathcal{I}(P) \):
\[
\mathcal{I}(P) = f(x_1^+) [x_1 - x_0] + f(x_2^+) [x_2 - x_1] + \cdots + f(x_n^+) [x_n - x_{n-1}]
\]

We claim that, for any reasonable\(^{17}\) function \( f(x) \), if you take any reasonable\(^{18}\) sequence of these approximations you always get the exactly the same limiting value. We define \( \int_a^b f(x) \, dx \) to be this limiting value.

Let’s be more precise. We can take the limit of these approximations in two equivalent ways. Above we did this by taking the number of subintervals \( n \) to infinity. When we did this, the width of all the subintervals went to zero. With the formulation we are now using, simply taking the number of subintervals to be very large does not imply that they will all shrink in size. We could have one very large subinterval and a large number of tiny ones. Thus we take the limit we need by taking the width of the subintervals to zero. So for any choice \( P \), we define
\[
M(P) = \max \{x_1 - x_0, x_2 - x_1, \ldots, x_n - x_{n-1}\}
\]
that is the maximum width of the subintervals used in the approximation determined by \( P \). By forcing the maximum width to go to zero, the widths of all the subintervals go to zero.

We then define the definite integral as the limit
\[
\int_a^b f(x) \, dx = \lim_{M(P) \to 0} \mathcal{I}(P).
\]

Of course, one is now left with the question of determining when the above limit exists. A proof of the very general conditions which guarantee existence of this limit is beyond the scope of this course, so we instead give a weaker result (with stronger conditions) which is far easier to prove.

For the rest of this section, assume

- that \( f(x) \) is continuous for \( a \leq x \leq b \),
- that \( f(x) \) is differentiable for \( a < x < b \), and

17 We’ll be more precise about what “reasonable” means shortly.
18 Again, we’ll explain this “reasonable” shortly.
that $f'(x)$ is bounded — ie $|f'(x)| \leq F$ for some constant $F$.

We will now show that, under these hypotheses, as $M(\mathbb{P})$ approaches zero, $\mathcal{I}(\mathbb{P})$ always approaches the area, $A$, between the graph of $y = f(x)$ and the $x$–axis, with $x$ running from $a$ to $b$.

These assumptions are chosen to make the argument particularly transparent. With a little more work one can weaken the hypotheses considerably. We are cheating a little by implicitly assuming that the area $A$ exists. In fact, one can adjust the argument below to remove this implicit assumption.

Consider $A_j$, the part of the area coming from $x_{j-1} \leq x \leq x_j$.

We have approximated this area by $f(x_j^*)[x_j - x_{j-1}]$ (see figure left).

Let $f(\overline{x}_j)$ and $f(x_j)$ be the largest and smallest values$^{19}$ of $f(x)$ for $x_{j-1} \leq x \leq x_j$. Then the true area is bounded by

$$f(x_j)[x_j - x_{j-1}] \leq A_j \leq f(\overline{x}_j)[x_j - x_{j-1}]$$

(see figure right).

Now since $f(\overline{x}_j) \leq f(x_j^*) \leq f(\overline{x}_j)$, we also know that

$$f(\overline{x}_j)[x_j - x_{j-1}] \leq f(x_j^*)[x_{j-1} - x_j] \leq f(\overline{x}_j)[x_j - x_{j-1}]$$

So both the true area, $A_j$, and our approximation of that area $f(x_j^*)[x_j - x_{j-1}]$ have to lie between $f(\overline{x}_j)[x_j - x_{j-1}]$ and $f(x_j)[x_j - x_{j-1}]$. Combining these bounds we have that the difference between the true area and our approximation of that area is bounded by

$$|A_j - f(x_j^*)[x_j - x_{j-1}]| \leq [f(\overline{x}_j) - f(x_j)] \cdot [x_j - x_{j-1}]$$

(To see this think about the smallest the true area can be and the largest our approximation can be and vice versa.)

$^{19}$ Here we are using the extreme value theorem — its proof is beyond the scope of this course. The theorem says that any continuous function on a closed interval must attain a minimum and maximum at least once. In this situation this implies that for any continuous function $f(x)$, there are $x_{j-1} \leq \overline{x}_j, \overline{x}_j \leq x_j$ such that $f(\overline{x}_j) \leq f(x) \leq f(\overline{x}_j)$ for all $x_{j-1} \leq x \leq x_j$. 

---

**INTEGRATION**

1.1 Definition of the Integral
Now since our function, \( f(x) \) is differentiable we can apply one of the main theorems we learned in mathematics 100 and 180 — the Mean Value Theorem\(^{20} \). The MVT implies that there exists a \( c \) between \( x_j \) and \( \bar{x}_j \) such that

\[
f(\bar{x}_j) - f(x_j) = f'(c) \cdot [\bar{x}_j - x_j]
\]

By the assumption that \( |f'(x)| \leq F \) for all \( x \) and the fact that \( x_j \) and \( \bar{x}_j \) must both be between \( x_{j-1} \) and \( x_j \)

\[
|f(\bar{x}_j) - f(x_j)| \leq F \cdot |\bar{x}_j - x_j| \leq F \cdot [x_j - x_{j-1}]
\]

Hence the error in this part of our approximation obeys

\[
|A_j - f(x_j^*)[x_j - x_{j-1}]| \leq F \cdot [x_j - x_{j-1}]^2.
\]

That was just the error in approximating \( A_j \). Now we bound the total error by combining the errors from approximating on all the subintervals. This gives

\[
|A - I(\mathcal{P})| = \left| \sum_{j=1}^{n} A_j - \sum_{j=1}^{n} f(x_j^*) [x_j - x_{j-1}] \right|
\]

\[
= \left| \sum_{j=1}^{n} \left( A_j - f(x_j^*) [x_j - x_{j-1}] \right) \right| \quad \text{triangle inequality}
\]

\[
\leq \sum_{j=1}^{n} \left| A_j - f(x_j^*) [x_j - x_{j-1}] \right|
\]

\[
\leq \sum_{j=1}^{n} F \cdot [x_j - x_{j-1}]^2 \quad \text{from above}
\]

Now do something a little sneaky. Replace one of these factors of \( [x_j - x_{j-1}] \) (which is just the width of the \( j^{th} \) subinterval) by the maximum width of the subintervals:

\[
\leq \sum_{j=1}^{n} F \cdot M(\mathcal{P}) \cdot [x_j - x_{j-1}] \quad F \text{ and } M(\mathcal{P}) \text{ are constant}
\]

\[
\leq F \cdot M(\mathcal{P}) \cdot \sum_{j=1}^{n} [x_j - x_{j-1}] \quad \text{sum is total width}
\]

\[
= F \cdot M(\mathcal{P}) \cdot (b - a).
\]

\(^{20}\) Recall that the mean value theorem states that for a function continuous on \( [a, b] \) and differentiable on \( (a, b) \), there exists a number \( c \) between \( a \) and \( b \) so that

\[
f'(c) = \frac{f(b) - f(a)}{b - a}.
\]
Since \(a, b\) and \(F\) are fixed, this tends to zero as the maximum rectangle width \(M(\mathcal{P})\) tends to zero.

Thus, we have proven

**Theorem 1.1.19.**

Assume that \(f(x)\) is continuous for \(a \leq x \leq b\), and is differentiable for all \(a < x < b\) with \(|f'(x)| \leq F\), for some constant \(F\). Then, as the maximum rectangle width \(M(\mathcal{P})\) tends to zero, \(I(\mathcal{P})\) always converges to \(A\), the area between the graph of \(y = f(x)\) and the \(x\)-axis, with \(x\) running from \(a\) to \(b\).

---

### 1.2 Basic properties of the definite integral

When we studied limits and derivatives, we developed methods for taking limits or derivatives of “complicated functions” like \(f(x) = x^2 + \sin(x)\) by understanding how limits and derivatives interact with basic arithmetic operations like addition and subtraction. This allowed us to reduce the problem into one of computing derivatives of simpler functions like \(x^2\) and \(\sin(x)\). Along the way we established simple rules such as

\[
\lim_{x \to a} (f(x) + g(x)) = \lim_{x \to a} f(x) + \lim_{x \to a} g(x) \quad \text{and} \quad \frac{d}{dx} (f(x) + g(x)) = \frac{df}{dx} + \frac{dg}{dx}
\]

Some of these rules have very natural analogues for integrals and we discuss them below. Unfortunately the analogous rules for integrals of products of functions or integrals of compositions of functions are more complicated than those for limits or derivatives. We discuss those rules at length in subsequent sections. For now let us consider some of the simpler rules of the arithmetic of integrals.
**Theorem 1.2.1 (Arithmetic of Integration).**

Let $a$, $b$ and $A$, $B$, $C$ be real numbers. Let the functions $f(x)$ and $g(x)$ be integrable on an interval that contains $a$ and $b$. Then

(a) \[ \int_a^b (f(x) + g(x)) \, dx = \int_a^b f(x) \, dx + \int_a^b g(x) \, dx \]

(b) \[ \int_a^b (f(x) - g(x)) \, dx = \int_a^b f(x) \, dx - \int_a^b g(x) \, dx \]

(c) \[ \int_a^b Cf(x) \, dx = C \cdot \int_a^b f(x) \, dx \]

Combining these three rules we have

(d) \[ \int_a^b (Af(x) + Bg(x)) \, dx = A \int_a^b f(x) \, dx + B \int_a^b g(x) \, dx \]

That is, integrals depend linearly on the integrand.

(e) \[ \int_a^b 1 \, dx = \int_a^b 1 \cdot dx = b - a \]

It is not too hard to prove this result from the definition of the definite integral. Additionally we only really need to prove (d) and (e) since

- (a) follows from (d) by setting $A = B = 1$,
- (b) follows from (d) by setting $A = 1, B = -1$, and
- (c) follows from (d) by setting $A = C, B = 0$.

**Proof.** As noted above, it suffices for us to prove (d) and (e). Since (e) is easier, we will start with that. It is also a good warm-up for (d).

- The definite integral in (e), $\int_a^b 1 \, dx$, can be interpreted geometrically as the area of the rectangle with height 1 running from $x = a$ to $x = b$; this area is clearly $b - a$. We can also prove this formula from the definition of the integral (Definition 1.1.9):

\[
\int_a^b 1 \, dx = \lim_{n \to \infty} \sum_{i=1}^{n} f(x_{i,n}) \frac{b-a}{n} \quad \text{by definition}
\]
\[
= \lim_{n \to \infty} \sum_{i=1}^{n} \frac{b-a}{n}
\]
\[
= \lim_{n \to \infty} (b-a) \sum_{i=1}^{n} \frac{1}{n}
\]
\[
= \lim_{n \to \infty} (b-a)
\]
\[
= b - a
\]
as required.

• To prove (d) let us start by defining \( h(x) = Af(x) + Bg(x) \) and then we need to express the integral of \( h(x) \) in terms of those of \( f(x) \) and \( g(x) \). We use Definition 1.1.9 and some algebraic manipulations\(^{21}\) to arrive at the result.

\[
\int_a^b h(x)\,dx = \sum_{i=1}^n h(x_{i,n}) \cdot \frac{b-a}{n}
\]

by Definition 1.1.9

\[
= \sum_{i=1}^n \left( Af(x_{i,n}) + Bg(x_{i,n}) \right) \cdot \frac{b-a}{n}
\]

\[
= \sum_{i=1}^n \left( Af(x_{i,n}) \cdot \frac{b-a}{n} + Bg(x_{i,n}) \cdot \frac{b-a}{n} \right)
\]

by Theorem 1.1.5(b)

\[
= \left( \sum_{i=1}^n Af(x_{i,n}) \cdot \frac{b-a}{n} \right) + \left( \sum_{i=1}^n Bg(x_{i,n}) \cdot \frac{b-a}{n} \right)
\]

by Theorem 1.1.5(a)

\[
= A \left( \sum_{i=1}^n f(x_{i,n}) \cdot \frac{b-a}{n} \right) + B \left( \sum_{i=1}^n g(x_{i,n}) \cdot \frac{b-a}{n} \right)
\]

as required.

Using this Theorem we can integrate sums, differences and constant multiples of functions we know how to integrate. For example:

**Example 1.2.2**

In Example 1.1.1 we saw that \( \int_0^1 e^x\,dx = e - 1 \). So

\[
\int_0^1 (e^x + 7)\,dx = \int_0^1 e^x\,dx + 7 \int_0^1 1\,dx
\]

by Theorem 1.2.1(d) with \( A = 1, f(x) = e^x, B = 7, g(x) = 1 \)

\[
= (e - 1) + 7 \times (1 - 0)
\]

by Example 1.1.1 and Theorem 1.2.1(e)

\[
= e + 6
\]

**Example 1.2.2**

When we gave the formal definition of \( \int_a^b f(x)\,dx \) in Definition 1.1.9 we explained that the integral could be interpreted as the signed area between the curve \( y = f(x) \) and the

\(^{21}\) Now is a good time to look back at Theorem 1.1.5.
INTEGRATION

1.2 BASIC PROPERTIES OF THE DEFINITE INTEGRAL

x-axis on the interval \([a, b]\). In order for this interpretation to make sense we required that \(a < b\), and though we remarked that the integral makes sense when \(a > b\) we did not explain any further. Thankfully there is an easy way to express the integral \(\int_a^b f(x)\,dx\) in terms of \(\int_b^a f(x)\,dx\) — making it always possible to write an integral so the lower limit of integration is less than the upper limit of integration. Theorem 1.2.3, below, tell us that, for example, \(\int_3^7 e^x\,dx = -\int_7^3 e^x\,dx\). The same theorem also provides us with two other simple manipulations of the limits of integration.

**Theorem 1.2.3 (Arithmetic for the Domain of Integration).**

Let \(a, b, c\) be real numbers. Let the function \(f(x)\) be integrable on an interval that contains \(a, b\) and \(c\). Then

(a) \(\int_a^a f(x)\,dx = 0\)

(b) \(\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx\)

(c) \(\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx\)

The proof of this statement is not too difficult.

**Proof.** Let us prove the statements in order.

- Consider the definition of the definite integral

\[
\int_a^b f(x)\,dx = \lim_{n \to \infty} \sum_{i=1}^{n} f(x_{i,n}^*) \cdot \frac{b-a}{n}
\]

If we now substitute \(b = a\) in this expression we have

\[
\int_a^a f(x)\,dx = \lim_{n \to \infty} \sum_{i=1}^{n} f(x_{i,n}^*) \cdot \frac{a-a}{n} = 0
\]

\[
= \lim_{n \to \infty} 0
\]

\[
= 0
\]

as required.

- Consider now the definite integral \(\int_a^b f(x)\,dx\). We will sneak up on the proof by first examining Riemann sum approximations to both this and \(\int_b^a f(x)\,dx\). The midpoint
Riemann sum approximation to \( \int_a^b f(x) \, dx \) with 4 subintervals (so that each subinterval has width \( \frac{b-a}{4} \)) is

\[
\left\{ f\left(a + \frac{1}{2} \frac{b-a}{4}\right) + f\left(a + \frac{3}{2} \frac{b-a}{4}\right) + f\left(a + \frac{5}{2} \frac{b-a}{4}\right) + f\left(a + \frac{7}{2} \frac{b-a}{4}\right) \right\} \cdot \frac{b-a}{4}
\]

Now we do the same for \( \int_b^a f(x) \, dx \) with 4 subintervals. Note that \( b \) is now the lower limit on the integral and \( a \) is now the upper limit on the integral. This is likely to cause confusion when we write out the Riemann sum, so we’ll temporarily rename \( b \) to \( A \) and \( a \) to \( B \). The midpoint Riemann sum approximation to \( \int_A^B f(x) \, dx \) with 4 subintervals is

\[
\left\{ f\left(A + \frac{1}{2} \frac{B-A}{4}\right) + f\left(A + \frac{3}{2} \frac{B-A}{4}\right) + f\left(A + \frac{5}{2} \frac{B-A}{4}\right) + f\left(A + \frac{7}{2} \frac{B-A}{4}\right) \right\} \cdot \frac{B-A}{4}
\]

Now recalling that \( A = b \) and \( B = a \), we have that the midpoint Riemann sum approximation to \( \int_b^a f(x) \, dx \) with 4 subintervals is

\[
\left\{ f\left(\frac{7}{8} a + \frac{1}{8} b\right) + f\left(\frac{5}{8} a + \frac{3}{8} b\right) + f\left(\frac{3}{8} a + \frac{5}{8} b\right) + f\left(\frac{1}{8} a + \frac{7}{8} b\right) \right\} \cdot \frac{a-b}{4}
\]

Thus we see that the Riemann sums for the two integrals are nearly identical — the only difference being the factor of \( \frac{b-a}{4} \) versus \( \frac{a-b}{4} \). Hence the two Riemann sums are negatives of each other.

The same computation with \( n \) subintervals shows that the midpoint Riemann sum approximations to \( \int_a^b f(x) \, dx \) and \( \int_b^a f(x) \, dx \) with \( n \) subintervals are negatives of each other. Taking the limit \( n \to \infty \) gives \( \int_a^b f(x) \, dx = -\int_a^b f(x) \, dx \).

- Finally consider (c) — we will not give a formal proof of this, but instead will interpret it geometrically. Indeed one can also interpret (a) geometrically. In both cases these become statements about areas:

\[
\int_a^a f(x) \, dx = 0 \quad \text{and} \quad \int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx
\]

are

\[
\text{Area}\{ (x, y) \mid a \leq x \leq a, \, 0 \leq y \leq f(x) \} = 0
\]

and

\[
\text{Area}\{ (x, y) \mid a \leq x \leq b, \, 0 \leq y \leq f(x) \} = \text{Area}\{ (x, y) \mid a \leq x \leq c, \, 0 \leq y \leq f(x) \} + \text{Area}\{ (x, y) \mid c \leq x \leq b, \, 0 \leq y \leq f(x) \}
\]

respectively. Both of these geometric statements are intuitively obvious. See the figures below.
Note that we have assumed that $a \leq c \leq b$ and that $f(x) \geq 0$. One can remove these restrictions and also make the proof more formal, but it becomes quite tedious and less intuitive.

Example 1.2.4

Back in Example 1.1.14 we saw that when $b > 0$ \( \int_{0}^{b} x \, dx = \frac{b^2}{2} \). We’ll now verify that \( \int_{0}^{b} x \, dx = \frac{b^2}{2} \) is still true when $b = 0$ and also when $b < 0$.

- First consider $b = 0$. Then the statement \( \int_{0}^{b} x \, dx = \frac{b^2}{2} \) becomes

\[
\int_{0}^{0} x \, dx = 0
\]

This is an immediate consequence of Theorem 1.2.3(a).

- Now consider $b < 0$. Let us write $B = -b$, so that $B > 0$. In Example 1.1.14 we saw that

\[
\int_{-B}^{0} x \, dx = -\frac{B^2}{2}.
\]

So we have

\[
\int_{0}^{b} x \, dx = \int_{0}^{-B} x \, dx = -\int_{-B}^{0} x \, dx \quad \text{by Theorem 1.2.3(b)}
\]

\[
= -\left( -\frac{B^2}{2} \right) \quad \text{by Example 1.1.14}
\]

\[
= \frac{B^2}{2} = \frac{b^2}{2}
\]

We have now shown that

\[
\int_{0}^{b} x \, dx = \frac{b^2}{2} \quad \text{for all real numbers } b
\]
Example 1.2.5

Applying Theorem 1.2.3 yet again, we have, for all real numbers \( a \) and \( b \),

\[
\int_a^b x \, dx = \int_0^a x \, dx + \int_a^b x \, dx \quad \text{by Theorem 1.2.3(c) with } c = 0
\]

\[
= \int_0^b x \, dx - \int_0^a x \, dx \quad \text{by Theorem 1.2.3(b)}
\]

\[
= \frac{b^2 - a^2}{2} \quad \text{by Example 1.2.4, twice}
\]

We can also understand this result geometrically.

- (left) When \( 0 < a < b \), the integral represents the area in green which is the difference of two right–angle triangles — the larger with area \( \frac{b^2}{2} \) and the smaller with area \( \frac{a^2}{2} \).

- (centre) When \( a < 0 < b \), the integral represents the signed area of the two displayed triangles. The one above the axis has area \( \frac{b^2}{2} \) while the one below has area \( -\frac{a^2}{2} \) (since it is below the axis).

- (right) When \( a < b < 0 \), the integral represents the signed area in purple of the difference between the two triangles — the larger with area \( -\frac{a^2}{2} \) and the smaller with area \( -\frac{b^2}{2} \).

Theorem 1.2.3(c) shows us how we can split an integral over a larger interval into one over two (or more) smaller intervals. This is particularly useful for dealing with piecewise functions, like \( |x| \).
Example 1.2.6

Using Theorem 1.2.3, we can readily evaluate integrals involving $|x|$. First, recall that

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

Now consider (for example) $\int_{-2}^{3} |x| \, dx$. Since the integrand changes at $x = 0$, it makes sense to split the interval of integration at that point:

$$
\int_{-2}^{3} |x| \, dx = \int_{-2}^{0} |x| \, dx + \int_{0}^{3} |x| \, dx 
$$

by Theorem 1.2.3

$$
= \int_{-2}^{0} (-x) \, dx + \int_{0}^{3} x \, dx 
$$

by definition of $|x|$

$$
= -\int_{-2}^{0} x \, dx + \int_{0}^{3} x \, dx 
$$

by Theorem 1.2.1(c)

$$
= -(-2^2/2) + (3^2/2) = (4 + 9)/2 = 13/2
$$

We can go further still — given a function $f(x)$ we can rewrite the integral of $f(|x|)$ in terms of the integral of $f(x)$ and $f(-x)$.

$$
\int_{-1}^{1} f(|x|) \, dx = \int_{-1}^{0} f(|x|) \, dx + \int_{0}^{1} f(|x|) \, dx 
$$

$$
= \int_{-1}^{0} f(-x) \, dx + \int_{0}^{1} f(x) \, dx 
$$

Example 1.2.6

Here is a more concrete example.

Example 1.2.7

Let us compute $\int_{-1}^{1} (1 - |x|) \, dx$ again. In Example 1.1.15 we evaluated this integral by interpreting it as the area of a triangle. This time we are going to use only the properties given in Theorems 1.2.1 and 1.2.3 and the facts that

$$
\int_{a}^{b} dx = b - a \quad \text{and} \quad \int_{a}^{b} x \, dx = \frac{b^2 - a^2}{2}
$$

That $\int_{a}^{b} dx = b - a$ is part (e) of Theorem 1.2.1. We saw that $\int_{a}^{b} x \, dx = \frac{b^2 - a^2}{2}$ in Example 1.2.5.

First we are going to get rid of the absolute value signs by splitting the interval over which we integrate. Recalling that $|x| = x$ whenever $x \geq 0$ and $|x| = -x$ whenever $x \leq 0$,
we split the interval by Theorem 1.2.3(c),

\[ \int_{-1}^{1} (1 - |x|) \, dx = \int_{-1}^{0} (1 - |x|) \, dx + \int_{0}^{1} (1 - |x|) \, dx \]

\[ = \int_{-1}^{0} (1 - (-x)) \, dx + \int_{0}^{1} (1 - x) \, dx \]

\[ = \int_{-1}^{0} (1 + x) \, dx + \int_{0}^{1} (1 - x) \, dx \]

Now we apply parts (a) and (b) of Theorem 1.2.1, and then

\[ \int_{-1}^{1} [1 - |x|] \, dx = \int_{-1}^{0} 1 \, dx + \int_{0}^{1} x \, dx + \int_{0}^{1} 1 \, dx - \int_{0}^{1} x \, dx \]

\[ = [0 - (-1)] + \frac{0^2 - (-1)^2}{2} + [1 - 0] - \frac{1^2 - 0^2}{2} \]

\[ = 1 \]

1.2.1 More properties of integration: even and odd functions

Recall\textsuperscript{22} the following definition

\begin{definition}

Let \( f(x) \) be a function. Then,

- we say that \( f(x) \) is even when \( f(x) = f(-x) \) for all \( x \), and
- we say that \( f(x) \) is odd when \( f(x) = -f(-x) \) for all \( x \).

\end{definition}

Of course most functions are neither even nor odd, but many of the standard functions you know are.

\begin{example}[Even functions]

- Three examples of even functions are \( f(x) = |x| \), \( f(x) = \cos x \) and \( f(x) = x^2 \). In fact, if \( f(x) \) is any even power of \( x \), then \( f(x) \) is an even function.

\end{example}

\textsuperscript{22} We haven’t done this in this course, but you should have seen it in your differential calculus course or perhaps even earlier.
• The part of the graph \( y = f(x) \) with \( x \leq 0 \), may be constructed by drawing the part of the graph with \( x \geq 0 \) (as in the figure on the left below) and then reflecting it in the \( y \)-axis (as in the figure on the right below).

\[
\begin{align*}
\int_{0}^{a} f(x) \, dx &= \int_{-a}^{0} f(x) \, dx
\end{align*}
\]

Example 1.2.9

Example 1.2.10 (Odd functions)

• Three examples of odd functions are \( f(x) = \sin x \), \( f(x) = \tan x \) and \( f(x) = x^3 \). In fact, if \( f(x) \) is any odd power of \( x \), then \( f(x) \) is an odd function.

• The part of the graph \( y = f(x) \) with \( x \leq 0 \), may be constructed by drawing the part of the graph with \( x \geq 0 \) (like the solid line in the figure on the left below) and then reflecting it in the \( y \)-axis (like the dashed line in the figure on the left below) and then reflecting the result in the \( x \)-axis (i.e. flipping it upside down, like in the figure on the right, below).
• In particular, if \( f(x) \) is an odd function and \( a > 0 \), then the signed areas of the two sets
\[
\begin{align*}
\{(x, y) \mid 0 \leq x \leq a \text{ and } y \text{ is between } 0 \text{ and } f(x)\} \\
\{(x, y) \mid -a \leq x \leq 0 \text{ and } y \text{ is between } 0 \text{ and } f(x)\}
\end{align*}
\]
are negatives of each other — to get from the first set to the second set, you flip it upside down, in addition to reflecting it in the \( x \)-axis. That is
\[
\int_{0}^{a} f(x) \, dx = -\int_{-a}^{0} f(x) \, dx
\]

We can exploit the symmetries noted in the examples above, namely
\[
\begin{align*}
\int_{0}^{a} f(x) \, dx &= \int_{-a}^{0} f(x) \, dx & \text{for } f \text{ even} \\
\int_{0}^{a} f(x) \, dx &= -\int_{-a}^{0} f(x) \, dx & \text{for } f \text{ odd}
\end{align*}
\]

Together with Theorem 1.2.3,
\[
\int_{-a}^{a} f(x) \, dx = \int_{-a}^{0} f(x) \, dx + \int_{0}^{a} f(x) \, dx
\]

in order to simplify the integration of even and odd functions over intervals of the form \([-a, a]\).

**Theorem 1.2.11 (Even and Odd).**

Let \( a > 0 \).

(a) If \( f(x) \) is an even function, then
\[
\int_{-a}^{a} f(x) \, dx = 2 \int_{0}^{a} f(x) \, dx
\]

(b) If \( f(x) \) is an odd function, then
\[
\int_{-a}^{a} f(x) \, dx = 0
\]

**Proof.** For any function
\[
\int_{-a}^{a} f(x) \, dx = \int_{-a}^{0} f(x) \, dx + \int_{0}^{a} f(x) \, dx
\]

When \( f \) is even, the two terms on the right hand side are equal. When \( f \) is odd, the two terms on the right hand side are negatives of each other. \(\square\)
1.2.2 Optional — More properties of integration: inequalities for integrals

We are still unable to integrate many functions, however with a little work we can infer bounds on integrals from bounds on their integrands.

Theorem 1.2.12 (Inequalities for Integrals).

Let \( a \leq b \) be real numbers and let the functions \( f(x) \) and \( g(x) \) be integrable on the interval \( a \leq x \leq b \).

(a) If \( f(x) \geq 0 \) for all \( a \leq x \leq b \), then
\[
\int_a^b f(x) \, dx \geq 0
\]

(b) If there are constants \( m \) and \( M \) such that \( m \leq f(x) \leq M \) for all \( a \leq x \leq b \), then
\[
m(b - a) \leq \int_a^b f(x) \, dx \leq M(b - a)
\]

(c) If \( f(x) \leq g(x) \) for all \( a \leq x \leq b \), then
\[
\int_a^b f(x) \, dx \leq \int_a^b g(x) \, dx
\]

(d) We have
\[
\left| \int_a^b f(x) \, dx \right| \leq \int_a^b |f(x)| \, dx
\]

Proof. (a) By interpreting the integral as the signed area, this statement simply says that if the curve \( y = f(x) \) lies above the \( x \)-axis and \( a \leq b \), then the signed area of \( \{(x, y) \mid a \leq x \leq b, 0 \leq y \leq f(x)\} \) is at least zero. This is quite clear. Alternatively, we could argue more algebraically from Definition 1.1.9. We observe that when we define \( \int_a^b f(x) \, dx \) via Riemann sums, every summand, \( f(x_{i,n}) \frac{b-a}{n} \geq 0 \). Thus the whole sum is nonnegative and consequently, so is the limit, and thus so is the integral.

(b) We can argue this from (a) with a little massaging. Let \( g(x) = M - f(x) \), then since \( f(x) \leq M \), we have \( g(x) = M - f(x) \geq 0 \) so that
\[
\int_a^b (M - f(x)) \, dx = \int_a^b g(x) \, dx \geq 0.
\]
but we also have
\[ \int_a^b (M - f(x)) \, dx = \int_a^b M \, dx - \int_a^b f(x) \, dx = M(b - a) - \int_a^b f(x) \, dx \]
Thus
\[ M(b - a) - \int_a^b f(x) \, dx \geq 0 \]
rearrange
\[ M(b - a) \geq \int_a^b f(x) \, dx \]
as required. The argument showing \( \int_a^b f(x) \, dx \geq m(b - a) \) is similar.

(c) Now let \( h(x) = g(x) - f(x) \). Since \( f(x) \leq g(x) \), we have \( h(x) = g(x) - f(x) \geq 0 \) so that
\[ \int_a^b (g(x) - f(x)) \, dx = \int_a^b h(x) \, dx \geq 0 \]
But we also have that
\[ \int_a^b (g(x) - f(x)) \, dx = \int_a^b g(x) \, dx - \int_a^b f(x) \, dx \]
Thus
\[ \int_a^b g(x) \, dx - \int_a^b f(x) \, dx \geq 0 \]
rearrange
\[ \int_a^b g(x) \, dx \geq \int_a^b f(x) \, dx \]
as required.

(d) For any \( x \), \(|f(x)|\) is either \( f(x) \) or \(-f(x)\) (depending on whether \( f(x) \) is positive or negative), so we certainly have
\[ f(x) \leq |f(x)| \quad \text{and} \quad -f(x) \leq |f(x)| \]
Applying part (c) to each of those inequalities gives
\[ \int_a^b f(x) \, dx \leq \int_a^b |f(x)| \, dx \quad \text{and} \quad -\int_a^b f(x) \, dx \leq \int_a^b |f(x)| \, dx \]
Now \(|\int_a^b f(x) \, dx|\) is either equal to \( \int_a^b f(x) \, dx \) or \(-\int_a^b f(x) \, dx \) (depending on whether the integral is positive or negative). In either case we can apply the above two inequalities to get the same result, namely
\[ \left| \int_a^b f(x) \, dx \right| \leq \int_a^b |f(x)| \, dx. \]
Example 1.2.13 \( \left( \int_0^{\pi/3} \sqrt{\cos x} \, dx \right) \)

Consider the integral

\[ \int_0^{\pi/3} \sqrt{\cos x} \, dx \]

This is not so easy to compute exactly\(^\text{23}\), but we can bound it quite quickly.

For \( x \) between 0 and \( \frac{\pi}{3} \), the function \( \cos x \) takes values\(^\text{24}\) between 1 and \( \frac{1}{2} \). Thus the function \( \sqrt{\cos x} \) takes values between 1 and \( \frac{1}{\sqrt{2}} \). That is

\[ \frac{1}{\sqrt{2}} \leq \sqrt{\cos x} \leq 1 \quad \text{for } 0 \leq x \leq \frac{\pi}{3}. \]

Consequently, by Theorem 1.2.12(b) with \( a = 0, b = \frac{\pi}{3}, m = \frac{1}{\sqrt{2}} \) and \( M = 1 \),

\[ \frac{\pi}{3\sqrt{2}} \leq \int_0^{\pi/3} \sqrt{\cos x} \, dx \leq \frac{\pi}{3} \]

Plugging these expressions into a calculator gives us

\[ 0.7404804898 \leq \int_0^{\pi/3} \sqrt{\cos x} \, dx \leq 1.047197551 \]

---

1.3 The Fundamental Theorem of Calculus

We have spent quite a few pages (and lectures) talking about definite integrals, what they are (Definition 1.1.9), when they exist (Theorem 1.1.10), how to compute some special cases (Section 1.1.4), some ways to manipulate them (Theorem 1.2.1 and 1.2.3) and how to bound them (Theorem 1.2.12). Conspicuously missing from all of this has been a discussion of how to compute them in general. It is high time we rectified that.

The single most important tool used to evaluate integrals is called “the fundamental theorem of calculus”. Its grand name is justified — it links the two branches of calculus by connecting derivatives to integrals. In so doing it also tells us how to compute integrals. Very roughly speaking the derivative of an integral is the original function. This fact allows us to compute integrals using antiderivatives\(^\text{25}\). Of course “very rough” is not enough — let’s be precise.

---

\( ^{23} \) It is not too hard to use Riemann sums and a computer to evaluate it numerically: 0.948025319 . . .

\( ^{24} \) You know the graphs of sine and cosine, so you should be able to work this out without too much difficulty.

\( ^{25} \) You learned these near the end of your differential calculus course. Now is a good time to revise — but we’ll go over them here since they are so important in what follows.
Theorem 1.3.1 (Fundamental Theorem of Calculus).

Let \( a < b \) and let \( f(x) \) be a function which is defined and continuous on \([a, b]\).

**Part 1:** Let \( F(x) = \int_a^x f(t) \, dt \) for any \( x \in [a, b] \). Then the function \( F(x) \) is differentiable and further

\[
F'(x) = f(x)
\]

**Part 2:** Let \( G(x) \) be any function which is defined and continuous on \([a, b]\). Further let \( G(x) \) be differentiable with \( G'(x) = f(x) \) for all \( a < x < b \). Then

\[
\int_a^b f(x) \, dx = G(b) - G(a) \quad \text{or equivalently} \quad \int_a^b G'(x) \, dx = G(b) - G(a)
\]

Before we prove this theorem and look at a bunch of examples of its application, it is important that we recall one definition from differential calculus — antiderivatives. If \( F'(x) = f(x) \) on some interval, then \( F(x) \) is called an antiderivative of \( f(x) \) on that interval. So Part 2 of the fundamental theorem of calculus tells us how to evaluate the definite integral of \( f(x) \) in terms of any of its antiderivatives — if \( G(x) \) is any antiderivative of \( f(x) \) then

\[
\int_a^b f(x) \, dx = G(b) - G(a)
\]

The form \( \int_a^b G'(x) \, dx = G(b) - G(a) \) of the fundamental theorem relates the rate of change of \( G(x) \) over the interval \( a \leq x \leq b \) to the net change of \( G \) between \( x = a \) and \( x = b \). For that reason, it is sometimes called the “net change theorem”.

We’ll start with a simple example. Then we’ll see why the fundamental theorem is true and then we’ll do many more, and more involved, examples.

Example 1.3.2 (A first example)

Consider the integral \( \int_a^b x \, dx \) which we have explored previously in Example 1.2.5.

- The integrand is \( f(x) = x \).
- We can readily verify that \( G(x) = \frac{x^2}{2} \) satisfies \( G'(x) = f(x) \) and so is an antiderivative of the integrand.
- Part 2 of Theorem 1.3.1 then tells us that

\[
\int_a^b f(x) \, dx = G(b) - G(a)
\]

\[
\int_a^b x \, dx = \frac{b^2}{2} - \frac{a^2}{2}
\]

which is precisely the result we obtained (with more work) in Example 1.2.5.
We do not give completely rigorous proofs of the two parts of the theorem — that is not really needed for this course. We just give the main ideas of the proofs so that you can understand why the theorem is true.

Part 1. We wish to show that if

$$F(x) = \int_a^x f(t) \, dt$$

then

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

- Assume that $F$ is the above integral and then consider $F'(x)$. By definition

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

- To understand this limit, we interpret the terms $F(x), F(x + h)$ as signed areas. To simplify this further, let’s only consider the case that $f$ is always nonnegative and that $h > 0$. These restrictions are not hard to remove, but the proof ideas are a bit cleaner if we keep them in place. Then we have

$$F(x + h) = \text{the area of the region } \{(t, y) \mid a \leq t \leq x + h, \ 0 \leq y \leq f(t) \}$$

$$F(x) = \text{the area of the region } \{(t, y) \mid a \leq t \leq x, \ 0 \leq y \leq f(t) \}$$

- Then the numerator

$$F(x + h) - F(x) = \text{the area of the region } \{(t, y) \mid x \leq t \leq x + h, \ 0 \leq y \leq f(t) \}$$

This is just the more darkly shaded region in the figure

- We will be taking the limit $h \to 0$. So suppose that $h$ is very small. Then, as $t$ runs from $x$ to $x = h$, $f(t)$ runs only over a very narrow range of values, all close to $f(x)$.26

- So the darkly shaded region is almost a rectangle of width $h$ and height $f(x)$ and so has an area which is very close to $f(x)h$. Thus $\frac{F(x + h) - F(x)}{h}$ is very close to $f(x)$.

26 Notice that if $f$ were discontinuous, then this might be false.
In the limit $h \to 0$, \( \frac{F(x+h) - F(x)}{h} \) becomes exactly $f(x)$, which is precisely what we want.

We can make the above more rigorous using the Mean Value Theorem.  

Part 2. We want to show that $\int_a^b f(t) \, dt = G(b) - G(a)$. To do this we exploit the fact that the derivative of a constant is zero.

Let 
\[
H(x) = \int_a^x f(t) \, dt - G(x) + G(a)
\]

Then the result we wish to prove is that $H(b) = 0$. We will do this by showing that $H(x) = 0$ for all $x$ between $a$ and $b$.

We first show that $H(x)$ is constant by computing its derivative:
\[
H'(x) = \frac{d}{dx} \int_a^x f(t) \, dt - \frac{d}{dx} (G(x)) + \frac{d}{dx} (G(a))
\]

Since $G(a)$ is a constant, its derivative is 0 and by assumption the derivative of $G(x)$ is just $f(x)$, so
\[
= \frac{d}{dx} \int_a^x f(t) \, dt - f(x)
\]

Now Part 1 of the theorem tells us that this derivative is just $f(x)$, so
\[
= f(x) - f(x) = 0
\]

Hence $H$ is constant.

To determine which constant we just compute $H(a)$:
\[
H(a) = \int_a^a f(t) \, dt - G(a) + G(a)
\]
\[
= \int_a^a f(t) \, dt \quad \text{by Theorem 1.2.3(a)}
\]
\[
= 0
\]

as required.

27 The MVT tells us that there is a number $c$ between $x + h$ and $x$ so that
\[
F'(c) = \frac{F(x+h) - F(x)}{(x+h) - x} = \frac{F(x+h) - F(x)}{h}
\]

But since $F'(x) = f(x)$, this tells us that
\[
\frac{F(x+h) - F(x)}{h} = f(c)
\]

where $c$ is trapped between $x + h$ and $x$. Now when we take the limit as $h \to 0$ we have that this number $c$ is squeezed to $x$ and the result follows.
The simple example we did above (Example 1.3.2), demonstrates the application of part 2 of the fundamental theorem of calculus. Before we do more examples (and there will be many more over the coming sections) we should do some examples illustrating the use of part 1 of the fundamental theorem of calculus. Then we’ll move on to part 2.

Example 1.3.3 \( \left( \frac{d}{dx} \int_0^x t \, dt \right) \)

Consider the integral \( \int_0^x t \, dt \). We know how to evaluate this — it is just Example 1.3.2 with \( a = 0, b = x \). So we have two ways to compute the derivative. We can evaluate the integral and then take the derivative, or we can apply Part 1 of the fundamental theorem. We’ll do both, and check that the two answers are the same.

First Example 1.3.2 gives

\[ F(x) = \int_0^x t \, dt = \frac{x^2}{2} \]

So of course \( F'(x) = x \). Second, Part 1 of the fundamental theorem of calculus tells us that the derivative of \( F(x) \) is just the integrand. That is, Part 1 of the fundamental theorem of calculus also gives \( F'(x) = x \).

Example 1.3.3

In the previous example we were able to evaluate the integral explicitly, so we did not need the fundamental theorem to determine its derivative. Here is an example that really does require the use of the fundamental theorem.

Example 1.3.4 \( \left( \frac{d}{dx} \int_0^x e^{-t^2} \, dt \right) \)

We would like to find \( \frac{d}{dx} \int_0^x e^{-t^2} \, dt \). In the previous example, we were able to compute the corresponding derivative in two ways — we could explicitly compute the integral and then differentiate the result, or we could apply part 1 of the fundamental theorem of calculus. In this example we do not know the integral explicitly. Indeed it is not possible to express \( \int_0^x e^{-t^2} \, dt \) as a finite combination of standard functions such as polynomials, exponentials, trigonometric functions and so on.

Despite this, we can find its derivative by just applying the first part of the fundamen-

\[ \int_0^x e^{-t^2} \, dt = x - \frac{x^3}{3 \cdot 1} + \frac{x^5}{5 \cdot 2} - \frac{x^7}{7 \cdot 3!} + \frac{x^9}{9 \cdot 4!} + \cdots + (-1)^k \frac{x^{2k+1}}{(2k+1) \cdot k!} + \cdots \]

But more on this in Chapter 3.
I
NTEGRATION

1.3 THE FUNDAMENTAL THEOREM OF CALCULUS

tal theorem of calculus with \( f(t) = e^{-t^2} \) and \( a = 0 \). That gives

\[
\frac{d}{dx} \int_0^x e^{-t^2} dt = \frac{d}{dx} \int_0^x f(t) dt
\]

\[
= f(x) = e^{-x^2}
\]

Example 1.3.4

Let us ratchet up the complexity of the previous example — we can make the limits of the integral more complicated functions. So consider the previous example with the upper limit \( x \) replaced by \( x^2 \):

Consider the integral \( \int_0^{x^2} e^{-t^2} dt \). We would like to compute its derivative with respect to \( x \) using part 1 of the fundamental theorem of calculus.

The fundamental theorem tells us how to compute the derivative of functions of the form \( \int_a^x f(t) dt \) but the integral at hand is not of the specified form because the upper limit we have is \( x^2 \), rather than \( x \), — so more care is required. Thankfully we can deal with this obstacle with only a little extra work. The trick is to define an auxiliary function by simply changing the upper limit to \( x \). That is, define

\[
E(x) = \int_0^x e^{-t^2} dt
\]

Then the integral we want to work with is

\[
E(x^2) = \int_0^{x^2} e^{-t^2} dt
\]

The derivative \( E'(x) \) can be found via part 1 of the fundamental theorem of calculus (as we did in Example 1.3.4) and is \( E'(x) = e^{-x^2} \). We can then use this fact with the chain rule to compute the derivative we need:

\[
\frac{d}{dx} \int_0^{x^2} e^{-t^2} dt = \frac{d}{dx} E(x^2)
\]

use the chain rule

\[
= 2x E'(x^2)
\]

\[
= 2xe^{-x^4}
\]

Example 1.3.5

What if both limits of integration are functions of \( x \)? We can still make this work, but we have to split the integral using Theorem 1.2.3.
Example 1.3.6 \( \frac{d}{dx} \int_x^{x^2} e^{-t^2} \, dt \)

Consider the integral

\[ \int_x^{x^2} e^{-t^2} \, dt \]

As was the case in the previous example, we have to do a little pre-processing before we can apply the fundamental theorem.

This time (by design), not only is the upper limit of integration \( x^2 \) rather than \( x \), but the lower limit of integration also depends on \( x \) — this is different from the integral \( \int_a^x f(t) \, dt \) in the fundamental theorem where the lower limit of integration is a constant.

Fortunately we can use the basic properties of integrals (Theorem 1.2.3(b) and (c)) to split \( \int_x^{x^2} e^{-t^2} \, dt \) into pieces whose derivatives we already know.

\[ \int_x^{x^2} e^{-t^2} \, dt = \int_x^0 e^{-t^2} \, dt + \int_0^{x^2} e^{-t^2} \, dt \]

by Theorem 1.2.3(c)

\[ = -\int_0^x e^{-t^2} \, dt + \int_0^{x^2} e^{-t^2} \, dt \]

by Theorem 1.2.3(b)

With this pre-processing, both integrals are of the right form. Using what we have learned in the previous two examples,

\[ \frac{d}{dx} \int_x^{x^2} e^{-t^2} \, dt = \frac{d}{dx} \left( -\int_0^x e^{-t^2} \, dt + \int_0^{x^2} e^{-t^2} \, dt \right) \]

\[ = -\frac{d}{dx} \int_0^x e^{-t^2} \, dt + \frac{d}{dx} \int_0^{x^2} e^{-t^2} \, dt \]

\[ = -e^{-x^2} + 2xe^{-x^4} \]

Example 1.3.6

Before we start to work with part 2 of the fundamental theorem, we need a little terminology and notation. First some terminology — you may have seen this definition in your differential calculus course.

**Definition 1.3.7 (Antiderivatives).**

Let \( f(x) \) and \( F(x) \) be functions. If \( F'(x) = f(x) \) on an interval, then we say that \( F(x) \) is an antiderivative of \( f(x) \) on that interval.

As we saw above, an antiderivative of \( f(x) = x \) is \( F(x) = x^2/2 \) — we can easily verify this by differentiation. Notice that \( x^2/2 + 3 \) is also an antiderivative of \( x \), as is \( x^2/2 + C \) for any constant \( C \). This observation gives us the following simple lemma.
Let $f(x)$ be a function and let $F(x)$ be an antiderivative of $f(x)$. Then $F(x) + C$ is also an antiderivative for any constant $C$. Further, every antiderivative of $f(x)$ must be of this form.

**Proof.** There are two parts to the lemma and we prove each in turn.

- Let $F(x)$ be an antiderivative of $f(x)$ and let $C$ be some constant. Then

$$\frac{d}{dx} (F(x) + C) = \frac{d}{dx} (F(x)) + \frac{d}{dx} (C) = f(x) + 0$$

since the derivative of a constant is zero, and by definition the derivative of $F(x)$ is just $f(x)$. Thus $F(x) + C$ is also an antiderivative of $f(x)$.

- Now let $F(x)$ and $G(x)$ both be antiderivatives of $f(x)$ — we will show that $G(x) = F(x) + C$ for some constant $C$. To do this let $H(x) = G(x) - F(x)$. Then

$$\frac{d}{dx} H(x) = \frac{d}{dx} (G(x) - F(x)) = \frac{d}{dx} G(x) - \frac{d}{dx} F(x) = f(x) - f(x) = 0$$

Since the derivative of $H(x)$ is zero, $H(x)$ must be a constant function\(^\text{29}\). Thus $H(x) = G(x) - F(x) = C$ for some constant $C$ and the result follows.

Based on the above lemma we have the following definition.

**Definition 1.3.9.**

The “indefinite integral of $f(x)$” is denoted by $\int f(x) \, dx$ and should be regarded as the general antiderivative of $f(x)$. In particular, if $F(x)$ is an antiderivative of $f(x)$ then

$$\int f(x) \, dx = F(x) + C$$

where the $C$ is an arbitrary constant. In this context, the constant $C$ is also often called a “constant of integration”.

\(^{29}\) This follows from the Mean Value Theorem. Say $H(x)$ were not constant, then there would be two numbers $a < b$ so that $H(a) \neq H(b)$. Then the MVT tells us that there is a number $c$ between $a$ and $b$ so that

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

Since both numerator and denominator are non-zero, we know the derivative at $c$ is nonzero. But this would contradict the assumption that derivative of $H$ is zero. Hence we cannot have $a < b$ with $H(a) \neq H(b)$ and so $H(x)$ must be constant.
Now we just need a tiny bit more notation.

**Notation 1.3.10.**

The symbol

\[ \int_a^b f(x) \, dx \]

denotes the change in an antiderivative of \( f(x) \) from \( x = a \) to \( x = b \). More precisely, let \( F(x) \) be any antiderivative of \( f(x) \). Then

\[ \int_a^b f(x) \, dx = F(x)|_a^b = F(b) - F(a) \]

Notice that this notation allows us to write part 2 of the fundamental theorem as

\[ \int_a^b f(x) \, dx = \int_a^b f(x) \, dx = F(x)|_a^b = F(b) - F(a) \]

Some texts also use an equivalent notation using square brackets:

\[ \int_a^b f(x) \, dx = \left[ F(x) \right]_a^b = F(b) - F(a). \]

You should be familiar with both notations.

We’ll soon develop some strategies for computing more complicated integrals. But for now, we’ll try a few integrals that are simple enough that we can just guess the answer. Of course, any antiderivative that we can guess we can also check — simply differentiate the guess and verify you get back to the original function:

\[ \frac{d}{dx} \int_a^b f(x) \, dx = f(x). \]

We do these examples in some detail to help us become comfortable finding indefinite integrals.

**Example 1.3.11**

Compute the definite integral \( \int_1^2 x \, dx \).

**Solution.** We have already seen, in Example 1.2.5, that \( \int_1^2 x \, dx = \frac{2^2 - 1^2}{2} = \frac{3}{2} \). We shall now rederive that result using the fundamental theorem of calculus.

- The main difficulty in this approach is finding the indefinite integral (an antiderivative) of \( x \). That is, we need to find a function \( F(x) \) whose derivative is \( x \). So think back to all the derivatives you computed last term and try to remember a function.

30 Of course, this assumes that you did your differential calculus course last term. If you did that course at a different time then please think back to that point in time. If it is long enough ago that you don’t quite remember when it was, then you should probably do some revision of derivatives of simple functions before proceeding further.
whose derivative was something like $x$.

- This shouldn’t be too hard — we recall that the derivatives of polynomials are polynomials. More precisely, we know that

$$\frac{d}{dx} x^n = nx^{n-1}$$

So if we want to end up with just $x = x^1$, we need to take $n = 2$. However this gives us

$$\frac{d}{dx} x^2 = 2x$$

- This is pretty close to what we want except for the factor of 2. Since this is a constant we can just divide both sides by 2 to obtain:

$$\frac{1}{2} \cdot \frac{d}{dx} x^2 = \frac{1}{2} \cdot 2x$$

which becomes

$$\frac{d}{dx} \frac{x^2}{2} = x$$

which is exactly what we need. It tells us that $x^2/2$ is an antiderivative of $x$.

- Once one has an antiderivative, it is easy to compute the indefinite integral

$$\int x \, dx = \frac{1}{2} x^2 + C$$

as well as the definite integral:

$$\int_1^2 x \, dx = \frac{1}{2} x^2 \bigg|_{1}^{2}$$

since $x^2/2$ is the antiderivative of $x$

$$= \frac{1}{2} (2^2 - 1^2) = \frac{3}{2}$$

While the previous example could be computed using signed areas, the following example would be very difficult to compute without using the fundamental theorem of calculus.

Compute $\int_0^{\pi/2} \sin x \, dx$.

**Solution.**

- Once again, the crux of the solution is guessing the antiderivative of $\sin x$ — that is finding a function whose derivative is $\sin x$. 

---

54
• The standard derivative that comes closest to \( \sin x \) is

\[
\frac{d}{dx} \cos x = -\sin x
\]

which is the derivative we want, multiplied by a factor of \(-1\).

• Just as we did in the previous example, we multiply this equation by a constant to remove this unwanted factor:

\[
(-1) \cdot \frac{d}{dx} \cos x = (-1) \cdot (-\sin x)
\]

\[
giving us \frac{d}{dx} (-\cos x) = \sin x
\]

This tells us that \(-\cos x\) is an antiderivative of \(\sin x\).

• Now it is straightforward to compute the integral:

\[
\int_{\pi/2}^{0} \sin x \, dx = -\cos x \bigg|_{\pi/2}^{0}
\]

since \(-\cos x\) is the antiderivative of \(\sin x\)

\[
= -\cos \frac{\pi}{2} + \cos 0
\]

\[
= 0 + 1 = 1
\]

Example 1.3.13

Find \( \int_{\frac{1}{2}}^{1} \frac{1}{x} \, dx \).

Solution.

• Once again, the crux of the solution is guessing a function whose derivative is \(\frac{1}{x}\). Our standard way to differentiate powers of \(x\), namely

\[
\frac{d}{dx} x^n = nx^{n-1},
\]

doesn’t work in this case — since it would require us to pick \(n = 0\) and this would give

\[
\frac{d}{dx} x^0 = \frac{d}{dx} 1 = 0.
\]
• Fortunately, we also know\(^{31}\) that
\[
\frac{d}{dx} \log x = \frac{1}{x}
\]
which is exactly the derivative we want.

• We’re now ready to compute the prescribed integral.
\[
\int_{1}^{2} \frac{1}{x} \, dx = \log x \bigg|_{1}^{2} \quad \text{since } \log x \text{ is an antiderivative of } \frac{1}{x}
\]
\[
= \log 2 - \log 1 \quad \text{since } \log 1 = 0
\]
\[
= \log 2
\]

Example 1.3.13

Example 1.3.14

Find \(\int_{-2}^{-1} \frac{1}{x} \, dx\).

Solution.

• As we saw in the last example,
\[
\frac{d}{dx} \log x = \frac{1}{x}
\]
and if we naively use this here, then we will obtain
\[
\int_{-2}^{-1} \frac{1}{x} \, dx = \log(-1) - \log(-2)
\]
which makes no sense since the logarithm is only defined for positive numbers\(^{32}\).

• We can work around this problem using a slight variation of the logarithm — \(\log |x|\).
  
  – When \(x > 0\), we know that \(|x| = x\) and so we have
  \[
  \log |x| = \log x \quad \text{differentiating gives us}
  \frac{d}{dx} \log |x| = \frac{d}{dx} \log x = \frac{1}{x}.
  \]

\(^{31}\) Recall that in most mathematics courses (especially this one) we use \(\log x\) without any indicated base to denote the natural logarithm — the logarithm base \(e\). Many widely used computer languages, like Java, C, Python, MATLAB, \(\cdots\), use \(\log(x)\) to denote the logarithm base \(e\) too. But many texts also use \(\ln x\) to denote the natural logarithm
\[
\log x = \log_e x = \ln x.
\]
The reader should be comfortable with all three notations for this function. They should also be aware that in different contexts — such as in chemistry or physics — it is common to use \(\log x\) to denote the logarithm base 10, while in computer science often \(\log x\) denotes the logarithm base 2. Context is key.

\(^{32}\) This is not entirely true — one can extend the definition of the logarithm to negative numbers, but to do so one needs to understand complex numbers which is a topic beyond the scope of this course.
- When \( x < 0 \) we have that \(|x| = -x\) and so

\[
\log |x| = \log(-x)
\]

differentiating with the chain rule gives

\[
\frac{d}{dx} \log |x| = \frac{d}{dx} \log(-x)
\]

\[
= \frac{1}{-x} \cdot (-1) = \frac{1}{x}
\]

- Indeed, more generally we should write the indefinite integral of \(1/\!x\) as

\[
\int \frac{1}{x} \, dx = \log |x| + C
\]

which is valid for all positive and negative \(x\). It is, however, undefined at \(x = 0\).

- We’re now ready to compute the prescribed integral.

\[
\int_{-2}^{-1} \frac{1}{x} \, dx = \log |x| \bigg|_{-2}^{-1}
\]

since \(\log |x|\) is an antiderivative of \(1/x\)

\[
= \log |-1| - \log |-2| = \log 1 - \log 2
\]

\[
= -\log 2 = \log 1/2.
\]

Example 1.3.14

This next example raises a nasty issue that requires a little care. We know that the function \(1/x\) is not defined at \(x = 0\) — so can we integrate over an interval that contains \(x = 0\) and still obtain an answer that makes sense? More generally can we integrate a function over an interval on which that function has discontinuities?

Example 1.3.15

Find \(\int_{-1}^{1} \frac{1}{x^2} \, dx\).

Solution. Beware that this is a particularly nasty example, which illustrates a booby trap hidden in the fundamental theorem of calculus. The booby trap explodes when the theorem is applied sloppily.

- The sloppy solution starts, as our previous examples have have, by finding an antiderivative of the integrand. In this case we know that

\[
\frac{d}{dx} \frac{1}{x} = -\frac{1}{x^2}
\]

which means that \(-x^{-1}\) is an antiderivative of \(x^{-2}\).

- This suggests (if we proceed naively) that

\[
\int_{-1}^{1} x^{-2} \, dx = -\frac{1}{x} \bigg|_{-1}^{1}
\]

since \(-1/x\) is an antiderivative of \(1/x^2\)

\[
= \frac{1}{1} - \left( -\frac{1}{-1} \right)
\]

\[
= -2
\]

Unfortunately,
• At this point we should really start to be concerned. This answer cannot be correct. Our integrand, being a square, is positive everywhere. So our integral represents the area of a region above the $x$–axis and must be positive.

• So what has gone wrong? The flaw in the computation is that the fundamental theorem of calculus, which says that

$$\text{if } F'(x) = f(x) \text{ then } \int_a^b f(x)\,dx = F(b) - F(a),$$

is only applicable when $F'(x)$ exists and equals $f(x)$ for all $x$ between $a$ and $b$.

• In this case $F'(x) = \frac{1}{x^2}$ does not exist for $x = 0$. So we cannot apply the fundamental theorem of calculus as we tried to above.

An integral, like $\int_1^1 \frac{1}{x^2}\,dx$, whose integrand is undefined somewhere in the domain of integration is called improper. We’ll give a more thorough treatment of improper integrals later in the text. For now, we’ll just say that the correct way to define (and evaluate) improper integrals is as a limit of well-defined approximating integrals. We shall later see that, not only is $\int_1^1 \frac{1}{x^2}\,dx$ not negative, it is infinite.

The above examples have illustrated how we can use the fundamental theorem of calculus to convert knowledge of derivatives into knowledge of integrals. We are now in a position to easily built a table of integrals. Here is a short table of the most important derivatives that we know.

<table>
<thead>
<tr>
<th>$F(x)$</th>
<th>$f(x) = F'(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1$</td>
<td>$0$</td>
</tr>
<tr>
<td>$x^n$</td>
<td>$nx^{n-1}$</td>
</tr>
<tr>
<td>$\sin x$</td>
<td>$\cos x$</td>
</tr>
<tr>
<td>$\cos x$</td>
<td>$-\sin x$</td>
</tr>
<tr>
<td>$\tan x$</td>
<td>$\sec^2 x$</td>
</tr>
<tr>
<td>$e^x$</td>
<td>$e^x$</td>
</tr>
<tr>
<td>$\log_e</td>
<td>x</td>
</tr>
<tr>
<td>$\arcsin x$</td>
<td>$\frac{1}{\sqrt{1-x^2}}$</td>
</tr>
<tr>
<td>$\arctan x$</td>
<td>$\frac{1}{1+x^2}$</td>
</tr>
</tbody>
</table>

Of course we know other derivatives, such as those of $\sec x$ and $\cot x$, however the ones listed above are arguably the most important ones. From this table (with a very little massaging) we can write down a short table of indefinite integrals.
Theorem 1.3.16 (Important indefinite integrals).

<table>
<thead>
<tr>
<th>$f(x)$</th>
<th>$\int f(x) , dx$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$x + C$</td>
</tr>
<tr>
<td>$x^n$</td>
<td>$\frac{1}{n+1} x^{n+1} + C$ provided that $n \neq -1$</td>
</tr>
<tr>
<td>$\frac{1}{x}$</td>
<td>$\log_e</td>
</tr>
<tr>
<td>$e^x$</td>
<td>$e^x + C$</td>
</tr>
<tr>
<td>$\sin x$</td>
<td>$- \cos x + C$</td>
</tr>
<tr>
<td>$\cos x$</td>
<td>$\sin x + C$</td>
</tr>
<tr>
<td>$\sec^2 x$</td>
<td>$\tan x + C$</td>
</tr>
<tr>
<td>$\frac{1}{\sqrt{1 - x^2}}$</td>
<td>$\arcsin x + C$</td>
</tr>
<tr>
<td>$\frac{1}{1 + x^2}$</td>
<td>$\arctan x + C$</td>
</tr>
</tbody>
</table>

Example 1.3.17

Find the following integrals

(i) $\int_2^7 e^x \, dx$

(ii) $\int_{-2}^2 \frac{1}{1 + x^2} \, dx$

(iii) $\int_0^3 (2x^3 + 7x - 2) \, dx$

Solution. We can proceed with each of these as before — find the antiderivative and then apply the fundamental theorem. The third integral is a little more complicated, but we can split it up into monomials using Theorem 1.2.1 and do each separately.
(i) An antiderivative of \(e^x\) is just \(e^x\), so
\[
\int_{2}^{7} e^x \, dx = e^x \bigg|_{2}^{7} = e^7 - e^2 = e^5(e^2 - 1).
\]

(ii) An antiderivative of \(\frac{1}{1+x^2}\) is \(\arctan(x)\), so
\[
\int_{-2}^{2} \frac{1}{1+x^2} \, dx = \arctan(x) \bigg|_{-2}^{2} = \arctan(2) - \arctan(-2)
\]
We can simplify this a little further by noting that \(\arctan(x)\) is an odd function, so \(\arctan(-2) = -\arctan(2)\) and thus our integral is
\[
= 2\arctan(2)
\]

(iii) We can proceed by splitting the integral using Theorem 1.2.1(d)
\[
\int_{0}^{3} (2x^3 + 7x - 2) \, dx = \int_{0}^{3} 2x^3 \, dx + \int_{0}^{3} 7x \, dx - \int_{0}^{3} 2 \, dx
\]
\[
= 2 \int_{0}^{3} x^3 \, dx + 7 \int_{0}^{3} x \, dx - 2 \int_{0}^{3} 1 \, dx
\]
and because we know that \(x^4/4, x^2/2, x\) are antiderivatives of \(x^3, x, 1\) respectively, this becomes
\[
= \left[ \frac{x^4}{2} \right]_{0}^{3} + \left[ \frac{7x^2}{2} \right]_{0}^{3} - [2x]_{0}^{3}
\]
\[
= \frac{81}{2} + \frac{7 \cdot 9}{2} - 6
\]
\[
= \frac{81 + 63 - 12}{2} = \frac{132}{2} = 66.
\]
We can also just find the antiderivative of the whole polynomial by finding the antiderivatives of each term of the polynomial and then recombining them. This is equivalent to what we have done above, but perhaps a little neater:
\[
\int_{0}^{3} (2x^3 + 7x - 2) \, dx = \left[ \frac{x^4}{2} + \frac{7x^2}{2} - 2x \right]_{0}^{3}
\]
\[
= \frac{81}{2} + \frac{7 \cdot 9}{2} - 6 = 66.
\]
1.4 Substitution

In the previous section we explored the fundamental theorem of calculus and the link it provides between definite integrals and antiderivatives. Indeed, integrals with simple integrands are usually evaluated via this link. In this section we start to explore methods for integrating more complicated integrals. We have already seen — via Theorem 1.2.1 — that integrals interact very nicely with addition, subtraction and multiplication by constants:

\[ \int_{a}^{b} (Af(x) +Bg(x)) \, dx = A \int_{a}^{b} f(x) \, dx + B \int_{a}^{b} g(x) \, dx \]

for \( A, B \) constants. By combining this with the list of indefinite integrals in Theorem 1.3.16, we can compute integrals of linear combinations of simple functions. For example

\[ \int_{1}^{4} (e^{x} - 2 \sin x + 3x^{2}) \, dx = \int_{1}^{4} e^{x} \, dx - 2 \int_{1}^{4} \sin x \, dx + 3 \int_{1}^{4} x^{2} \, dx \]

\[ = \left( e^{x} + (-2) \cdot (- \cos x) + 3 \frac{x^{3}}{3} \right) \bigg|_{1}^{4} \quad \text{and so on} \]

Of course there are a great many functions that can be approached in this way, however there are some very simple examples that cannot.

\[ \int \sin(\pi x) \, dx \quad \int xe^{x} \, dx \quad \int \frac{x}{x^{2} - 5x + 6} \, dx \]

In each case the integrands are not linear combinations of simpler functions; in order to compute them we need to understand how integrals (and antiderivatives) interact with compositions, products and quotients. We reached a very similar point in our differential calculus course where we understood the linearity of the derivative,

\[ \frac{d}{dx} (Af(x) + Bg(x)) = Af'(x) + Bg'(x) \]

but had not yet seen the chain, product and quotient rules\(^{33}\). While we will develop tools to find the second and third integrals in later sections, we should really start with how to integrate compositions of functions.

It is important to state up front, that in general one cannot write down the integral of the composition of two functions — even if those functions are simple. This is not because the integral does not exist. Rather it is because the integral cannot be written down as a finite combination of the standard functions we know. A very good example of this, which we encountered in Example 1.3.4, is the composition of \( e^{x} \) and \(-x^{2}\). Even though we know

\[ \int e^{x} \, dx = e^{x} + C \quad \text{and} \quad \int -x^{2} \, dx = -\frac{1}{3} x^{3} + C \]

33 If your memory of these rules is a little hazy then you really should go back and revise them before proceeding. You will definitely need a good grasp of the chain rule for what follows in this section.
there is no simple function that is equal to the indefinite integral

$$\int e^{-x^2} \, dx.$$  

even though the indefinite integral exists. In this way integration is very different from differentiation.

With that caveat out of the way, we can introduce the substitution rule. The substitution rule is obtained by antidifferentiating the chain rule. In some sense it is the chain rule in reverse. For completeness, let us restate the chain rule:

**Theorem 1.4.1 (The chain rule).**

Let $F(u)$ and $u(x)$ be differentiable functions and form their composition $F(u(x))$. Then

$$\frac{d}{dx} F(u(x)) = F'(u(x)) \cdot u'(x)$$

Equivalently, if $y(x) = F(u(x))$, then

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

Consider a function $f(u)$, which has antiderivative $F(u)$. Then we know that

$$\int f(u) \, du = \int F'(u) \, du = F(u) + C$$

Now take the above equation and substitute into it $u = u(x)$ — i.e. replace the variable $u$ with any (differentiable) function of $x$ to get

$$\int f(u) \, du \bigg|_{u=u(x)} = F(u(x)) + C$$

But now the right-hand side is a function of $x$, so we can differentiate it with respect to $x$ to get

$$\frac{d}{dx} F(u(x)) = F'(u(x)) \cdot u'(x)$$

This tells us that $F(u(x))$ is an antiderivative of the function $F'(u(x)) \cdot u'(x) = f(u(x))u'(x)$. Thus we know

$$\int f(u(x)) \cdot u'(x) \, dx = F(u(x)) + C = \int f(u) \, du \bigg|_{u=u(x)}$$

This is the substitution rule for indefinite integrals.
**Theorem 1.4.2** (The substitution rule — indefinite integral version).

For any differentiable function \( u(x) \):

\[
\int f(u(x))u'(x)\,dx = \int f(u)\,du \bigg|_{u=u(x)}
\]

In order to apply the substitution rule successfully we will have to write the integrand in the form \( f(u(x)) \cdot u'(x) \). To do this we need to make a good choice of the function \( u(x) \); after that it is not hard to then find \( f(u) \) and \( u'(x) \). Unfortunately there is no one strategy for choosing \( u(x) \). This can make applying the substitution rule more art than science. Here we suggest two possible strategies for picking \( u(x) \):

1. Factor the integrand and choose one of the factors to be \( u'(x) \). For this to work, you must be able to easily find the antiderivative of the chosen factor. The antiderivative will be \( u(x) \).

2. Look for a factor in the integrand that is a function with an argument that is more complicated than just “\( x \)”. That factor will play the role of \( f(u(x)) \). Choose \( u(x) \) to be the complicated argument.

Here are two examples which illustrate each of those strategies in turn.

**Example 1.4.3**

Consider the integral

\[
\int 9 \sin^8(x) \cos(x)\,dx
\]

We want to massage this into the form of the integrand in the substitution rule — namely \( f(u(x)) \cdot u'(x) \). Our integrand can be written as the product of the two factors

\[
\underbrace{9 \sin^8(x)}_{\text{first factor}} \cdot \underbrace{\cos(x)}_{\text{second factor}}
\]

and we start by determining (or guessing) which factor plays the role of \( u'(x) \). We can choose \( u'(x) = 9 \sin^8(x) \) or \( u'(x) = \cos(x) \).

- If we choose \( u'(x) = 9 \sin^8(x) \), then antidifferentiating this to find \( u(x) \) is really not very easy. So it is perhaps better to investigate the other choice before proceeding further with this one.

- If we choose \( u'(x) = \cos(x) \), then we know (Theorem 1.3.16) that \( u(x) = \sin(x) \). This also works nicely because it makes the other factor simplify quite a bit: \( 9 \sin^8(x) = 9u^8 \). This looks like the right way to go.

34 Thankfully this does become easier with experience and we recommend that the reader read some examples and then practice a LOT.
So we go with the second choice. Set $u'(x) = \cos(x), u(x) = \sin(x)$, then

$$
\int 9\sin^8(x) \cos(x) \, dx = \int 9u(x)^8 \cdot u'(x) \, dx
$$

$$
= \int 9u^8 \, du \bigg|_{u=\sin(x)} \quad \text{by the substitution rule}
$$

We are now left with the problem of antidifferentiating a monomial; this we can do with Theorem 1.3.16.

$$
= \left( u^9 + C \right) \bigg|_{u=\sin(x)}
$$

$$
= \sin^9(x) + C
$$

Note that $9\sin^8(x) \cos(x)$ is a function of $x$. So our answer, which is the indefinite integral of $9\sin^8(x) \cos(x)$, must also be a function of $x$. This is why we have substituted $u = \sin(x)$ in the last step of our solution — it makes our solution a function of $x$.

---

**Example 1.4.3**

Evaluate the integral

$$
\int 3x^2 \cos(x^3) \, dx
$$

*Solution.* Again we are going to use the substitution rule and helpfully our integrand is a product of two factors

$$
\frac{3x^2}{\text{first factor}} \cdot \frac{\cos(x^3)}{\text{second factor}}
$$

The second factor, $\cos(x^3)$ is a function, namely $\cos$, with a complicated argument, namely $x^3$. So we try $u(x) = x^3$. Then $u'(x) = 3x^2$, which is the other factor in the integrand. So the integral becomes

$$
\int 3x^2 \cos(x^3) \, dx = \int u'(x) \cos(u(x)) \, dx \quad \text{just swap order of factors}
$$

$$
= \int \cos(u(x)) \, du \bigg|_{u=x^3} \quad \text{by the substitution rule}
$$

$$
= \cos(u) \bigg|_{u=x^3}
$$

$$
= \sin(x^3) + C
$$

---

**Example 1.4.4**

Evaluate the integral

$$
\int 3x^2 \cos(x^3) \, dx
$$

*Solution.* Again we are going to use the substitution rule and helpfully our integrand is a product of two factors

$$
\frac{3x^2}{\text{first factor}} \cdot \frac{\cos(x^3)}{\text{second factor}}
$$

The second factor, $\cos(x^3)$ is a function, namely $\cos$, with a complicated argument, namely $x^3$. So we try $u(x) = x^3$. Then $u'(x) = 3x^2$, which is the other factor in the integrand. So the integral becomes

$$
\int 3x^2 \cos(x^3) \, dx = \int u'(x) \cos(u(x)) \, dx \quad \text{just swap order of factors}
$$

$$
= \int \cos(u(x)) \, du \bigg|_{u=x^3} \quad \text{by the substitution rule}
$$

$$
= \cos(u) \bigg|_{u=x^3}
$$

$$
= \sin(x^3) + C
$$
One more — we’ll use this to show how to use the substitution rule with definite integrals.

Example 1.4.5 \( \int_{0}^{1} e^{x} \sin(e^{x}) \, dx \)

Compute \( \int_{0}^{1} e^{x} \sin(e^{x}) \, dx \).

Solution. Again we use the substitution rule.

- The integrand is again the product of two factors and we can choose \( u'(x) = e^{x} \) or \( u'(x) = \sin(e^{x}) \).

- If we choose \( u'(x) = e^{x} \) then \( u(x) = e^{x} \) and the other factor becomes \( \sin(u) \) — this looks promising. Notice that if we applied the other strategy of looking for a complicated argument then we would arrive at the same choice.

- So we try \( u'(x) = e^{x} \) and \( u(x) = e^{x} \). This gives (if we ignore the limits of integration for a moment)

\[
\int e^{x} \sin(e^{x}) \, dx = \int \sin(u(x)) u'(x) \, dx \\
= \int \sin(u) \, du \bigg|_{u=e^{x}} \\
= (-\cos(u) + C) \bigg|_{u=e^{x}} \\
= -\cos(e^{x}) + C
\]

- But what happened to the limits of integration? We can incorporate them now. We have just shown that the indefinite integral is \( -\cos(e^{x}) \), so by the fundamental theorem of calculus

\[
\int_{0}^{1} e^{x} \sin(e^{x}) \, dx = [-\cos(e^{x})]_{0}^{1} \\
= -\cos(e^{1}) - (-\cos(e^{0})) \\
= -\cos(e) + \cos(1)
\]

Theorem 1.4.2, the substitution rule for indefinite integrals, tells us that if \( F(u) \) is any antiderivative for \( f(u) \), then \( F(u(x)) \) is an antiderivative for \( f(u(x))u'(x) \). So the funda-
mental theorem of calculus gives us
\[ \int_a^b f(u(x))u'(x) \, dx = \left. F(u(x)) \right|_{x=a}^{x=b} = F(u(b)) - F(u(a)) = \int_{u(a)}^{u(b)} f(u) \, du \]
since \( F(u) \) is an antiderivative for \( f(u) \)
and we have just found
\[
\textbf{Theorem 1.4.6 (The substitution rule — definite integral version).}
\]
For any differentiable function \( u(x) \):
\[
\int_a^b f(u(x))u'(x) \, dx = \int_{u(a)}^{u(b)} f(u) \, du
\]
Notice that to get from the integral on the left hand side to the integral on the right hand side you
- substitute\(^{35}\) \( u(x) \to u \) and \( u'(x) \, dx \to du \),
- set the lower limit for the \( u \) integral to the value of \( u \) (namely \( u(a) \)) that corresponds to the lower limit of the \( x \) integral (namely \( x = a \)), and
- set the upper limit for the \( u \) integral to the value of \( u \) (namely \( u(b) \)) that corresponds to the upper limit of the \( x \) integral (namely \( x = b \)).

Also note that we now have two ways to evaluate definite integrals of the form \( \int_a^b f(u(x))u'(x) \, dx \).
- We can find the indefinite integral \( \int f(u(x))u'(x) \, dx \), using Theorem 1.4.2, and then evaluate the result between \( x = a \) and \( x = b \). This is what was done in Example 1.4.5.
- Or we can apply Theorem 1.4.2. This entails finding the indefinite integral \( \int f(u) \, du \) and evaluating the result between \( u = u(a) \) and \( u = u(b) \). This is what we will do in the following example.

**Example 1.4.7** \( \int_0^1 x^2 \sin(x^3 + 1) \, dx \)

Compute
\[
\int_0^1 x^2 \sin(x^3 + 1) \, dx
\]

**Solution.**

\(^{35}\) A good way to remember this last step is that we replace \( \frac{du}{dx} \, dx \) by just \( du \) — which looks like we cancelled out the \( dx \) terms: \( \frac{du}{dx} \, dx = du \). While using “cancel the \( dx \)” is a good mnemonic (memory aid), you should not think of the derivative \( \frac{du}{dx} \) as a fraction — you are not dividing \( du \) by \( dx \).
In this example the integrand is already neatly factored into two pieces. While we could deploy either of our two strategies, it is perhaps easier in this case to choose $u(x)$ by looking for a complicated argument.

The second factor of the integrand is $\sin \left( x^3 + 1 \right)$, which is the function $\sin$ evaluated at $x^3 + 1$. So set $u(x) = x^3 + 1$, giving $u'(x) = 3x^2$ and $f(u) = \sin(u)$.

The first factor of the integrand is $x^2$ which is not quite $u^{-1}(x)$, however we can easily massage the integrand into the required form by multiplying and dividing by 3:

$$x^2 \sin \left( x^3 + 1 \right) = \frac{1}{3} \cdot 3x^2 \cdot \sin \left( x^3 + 1 \right).$$

We want this in the form of the substitution rule, so we do a little massaging:

$$\int_0^1 x^2 \sin \left( x^3 + 1 \right) dx = \int_0^1 \frac{1}{3} \cdot 3x^2 \cdot \sin \left( x^3 + 1 \right) dx$$

$$= \frac{1}{3} \int_0^1 \sin \left( x^3 + 1 \right) \cdot 3x^2 dx \quad \text{by Theorem 1.2.1(c)}$$

Now we are ready for the substitution rule:

$$\frac{1}{3} \int_0^1 \sin \left( x^3 + 1 \right) \cdot 3x^2 dx = \frac{1}{3} \int_0^1 \sin \left( x^3 + 1 \right) \cdot \frac{3x^2}{u'(x)} \cdot u'(x) dx$$

$$= \frac{1}{3} \int_0^1 f(u(x))u'(x) dx \quad \text{with } u(x) = x^3 + 1 \text{ and } f(u) = \sin(u)$$

$$= \frac{1}{3} \int_{u(0)}^{u(1)} f(u) du \quad \text{by the substitution rule}$$

$$= \frac{1}{3} \int_1^2 \sin(u) du \quad \text{since } u(0) = 1 \text{ and } u(1) = 2$$

$$= \frac{1}{3} \left[ -\cos(u) \right]_1^2$$

$$= \frac{1}{3} \left( -\cos(2) - (-\cos(1)) \right)$$

$$= \frac{\cos(1) - \cos(2)}{3}.$$ 

Example 1.4.7

There is another, and perhaps easier, way to view the manipulations in the previous example. Once you have chosen $u(x)$ you

- make the substitution $u(x) \rightarrow u$,

- replace $dx \rightarrow \frac{1}{u'(x)} du$. 
In so doing, we take the integral
\[
\int_a^b f(u(x)) \cdot u'(x) \, dx = \int_{u(a)}^{u(b)} f(u) \cdot u'(x) \cdot \frac{1}{u'(x)} \, du
\]
\[
= \int_{u(a)}^{u(b)} f(u) \, du
\]
which is exactly the substitution rule.

but we do not have to manipulate the integrand so as to make \( u'(x) \) explicit. Let us redo the previous example by this approach.

**Example 1.4.8 (Example 1.4.7 revisited)**

Compute the integral
\[
\int_0^1 x^2 \sin (x^3 + 1) \, dx
\]

**Solution.**

- We have already observed that one factor of the integrand is \( \sin (x^3 + 1) \), which is \( \sin \) evaluated at \( x^3 + 1 \). Thus we try setting \( u(x) = x^3 + 1 \).
- This makes \( u'(x) = 3x^2 \), and we replace \( u(x) = x^3 + 1 \rightarrow u \) and \( dx \rightarrow \frac{1}{u'(x)} \, du = \frac{1}{3x^2} \, du \):

\[
\int_0^1 x^2 \sin (x^3 + 1) \, dx = \int_{u(0)}^{u(1)} x^2 \sin (x^3 + 1) \cdot \frac{1}{3x^2} \, du
\]

\[
= \int_1^2 \sin(u) \cdot \frac{x^2}{3x^2} \, du
\]

\[
= \int_1^2 \frac{1}{3} \sin(u) \, du
\]

\[
= \frac{1}{3} \int_1^2 \sin(u) \, du
\]

which is precisely the integral we found in Example 1.4.7.

---

**Example 1.4.9**

Compute the indefinite integrals
\[
\int \sqrt{2x + 1} \, dx \quad \text{and} \quad \int e^{3x-2} \, dx
\]

**Solution.**
• Starting with the first integral, we see that it is not too hard to spot the complicated argument. If we set \( u(x) = 2x + 1 \) then the integrand is just \( \sqrt{u} \).

• Hence we substitute \( 2x + 1 \to u \) and \( dx \to \frac{1}{u'(x)}du = \frac{1}{2}du \):

\[
\int \sqrt{2x + 1} dx = \int \sqrt{u} \frac{1}{2} du = \int u^{1/2} \frac{1}{2} du = \left( \frac{2}{3}u^{3/2} \cdot \frac{1}{2} + C \right)_{u=2x+1} = \frac{1}{3}(2x + 1)^{3/2} + C
\]

• We can evaluate the second integral in much the same way. Set \( u(x) = 3x - 2 \) and replace \( dx \) by \( \frac{1}{u'(x)}du = \frac{1}{3}du \):

\[
\int e^{3x-2} dx = \int e^u \frac{1}{3} du = \left( \frac{1}{3}e^u + C \right)_{u=3x-2} = \frac{1}{3}e^{3x-2} + C
\]

This last example illustrates that substitution can be used to easily deal with arguments of the form \( ax + b \), i.e. that are linear functions of \( x \), and suggests the following theorem.

**Theorem 1.4.10.**

Let \( F(u) \) be an antiderivative of \( f(u) \) and let \( a, b \) be constants. Then

\[
\int f(ax + b) dx = \frac{1}{a} F(ax + b) + C
\]

**Proof.** We can show this using the substitution rule. Let \( u(x) = ax + b \) so \( u'(x) = a \), then

\[
\int f(ax + b) dx = \int f(u) \cdot \frac{1}{u'(x)} du = \int \frac{1}{a} f(u) du = \frac{1}{a} \int f(u) du = \frac{1}{a} F(u)_{u=ax+b} + C \text{ since } F(u) \text{ is an antiderivative of } f(u) = \frac{1}{a} F(ax + b) + C.
\]
Now we can do the following example using the substitution rule or the above theorem:

**Example 1.4.11** $\int_{0}^{\pi/2} \cos(3x) \, dx$

Compute $\int_{0}^{\pi/2} \cos(3x) \, dx$.

- In this example we should set $u = 3x$, and substitute $dx \rightarrow \frac{1}{u'(x)} \, du = \frac{1}{3} \, du$. When we do this we also have to convert the limits of the integral: $u(0) = 0$ and $u(\pi/2) = 3\pi/2$. This gives

  \[
  \int_{0}^{\pi/2} \cos(3x) \, dx = \int_{0}^{3\pi/2} \cos(u) \frac{1}{3} \, du \\
  = \left[ \frac{1}{3} \sin(u) \right]_{0}^{3\pi/2} \\
  = \sin(3\pi/2) - \sin(0) \\
  = -1 - 0 = -\frac{1}{3}.
  \]

- We can also do this example more directly using the above theorem. Since $\sin(x)$ is an antiderivative of $\cos(x)$, Theorem 1.4.10 tells us that $\frac{\sin(3x)}{3}$ is an antiderivative of $\cos(3x)$. Hence

  \[
  \int_{0}^{\pi/2} \cos(3x) \, dx = \left[ \frac{\sin(3x)}{3} \right]_{0}^{\pi/2} \\
  = \frac{\sin(3\pi/2) - \sin(0)}{3} \\
  = -\frac{1}{3}.
  \]

The rest of this section is just more examples of the substitution rule. We recommend that you after reading these that you practice many examples by yourself under exam conditions.

**Example 1.4.12** $\int_{0}^{1} x^2 \sin(1 - x^3) \, dx$

This integral looks a lot like that of Example 1.4.7. It makes sense to try $u(x) = 1 - x^3$ since it is the argument of $\sin(1 - x^3)$. We

- substitute $u = 1 - x^3$ and
- replace $dx$ with $\frac{1}{u'(x)} \, du = \frac{1}{-3x^2} \, du$,
when \( x = 0 \), we have \( u = 1 - 0^3 = 1 \) and

- when \( x = 1 \), we have \( u = 1 - 1^3 = 0 \).

So

\[
\int_0^1 x^2 \sin (1 - x^3) \, dx = \int_1^0 x^2 \sin(u) \cdot \frac{1}{-3x^2} \, du
\]

\[
= \int_1^0 -\frac{1}{3} \sin(u) \, du.
\]

Note that the lower limit of the \( u \)-integral, namely 1, is larger than the upper limit, which is 0. There is absolutely nothing wrong with that. We can simply evaluate the \( u \)-integral in the normal way. Since \( -\cos(u) \) is an antiderivative of \( \sin(u) \):

\[
= \left[ \frac{\cos(u)}{3} \right]^0_1
\]

\[
= \frac{\cos(0) - \cos(1)}{3}
\]

\[
= \frac{1 - \cos(1)}{3}.
\]

---

Example 1.4.13

Compute \( \int_0^1 \frac{1}{(2x+1)^3} \, dx \).

We could do this one using Theorem 1.4.10, but it's not too hard to do without. We can think of the integrand as the function “one over a cube” with the argument \( 2x + 1 \). So it makes sense to substitute \( u = 2x + 1 \). That is

- set \( u = 2x + 1 \) and

- replace \( dx \rightarrow \frac{1}{u'(x)} \, du = \frac{1}{2} \, du \).

- When \( x = 0 \), we have \( u = 2 \times 0 + 1 = 1 \) and

- when \( x = 1 \), we have \( u = 2 \times 1 + 1 = 3 \).
So
\[ \int_0^1 \frac{1}{(2x+1)^3} \, dx = \int_1^3 \frac{1}{u^3} \cdot \frac{1}{2} \, du \]
\[ = \frac{1}{2} \int_1^3 u^{-3} \, du \]
\[ = \frac{1}{2} \left[ \frac{u^{-2}}{-2} \right]_1^3 \]
\[ = \frac{1}{2} \left( \frac{1}{-2} - \frac{1}{9} - \frac{1}{-2} \cdot \frac{1}{1} \right) \]
\[ = \frac{1}{2} \left( \frac{1}{2} - \frac{1}{18} \right) = \frac{1}{2} \cdot \frac{8}{18} \]
\[ = \frac{2}{9} \]

**Example 1.4.13**

Evaluate \( \int_0^1 \frac{x}{1+x^2} \, dx \).

**Solution.**

- The integrand can be rewritten as \( x \cdot \frac{1}{1+x^2} \). This second factor suggests that we should try setting \( u = 1 + x^2 \) — and so we interpret the second factor as the function “one over” evaluated at argument \( 1 + x^2 \).

- With this choice we
  - set \( u = 1 + x^2 \),
  - substitute \( dx \to \frac{1}{2x} \, du \), and
  - translate the limits of integration: when \( x = 0 \), we have \( u = 1 + 0^2 = 1 \) and when \( x = 1 \), we have \( u = 1 + 1^2 = 2 \).

- The integral then becomes
  \[ \int_0^1 \frac{x}{1+x^2} \, dx = \int_1^2 \frac{x}{u} \cdot \frac{1}{2x} \, du \]
  \[ = \int_1^2 \frac{1}{2u} \, du \]
  \[ = \frac{1}{2} \left[ \log |u| \right]_1^2 \]
  \[ = \frac{\log 2 - \log 1}{2} = \frac{\log 2}{2} \].
Remember that we are using the notation “log” for the natural logarithm, i.e. the logarithm with base \(e\). You might also see it written as “\(\ln x\)”, or with the base made explicit as “\(\log_e x\)”. 

Example 1.4.14 

\[ \int x^3 \cos (x^4 + 2) \, dx \]

Compute the integral \( \int x^3 \cos (x^4 + 2) \, dx \).

Solution.

- The integrand is the product of cos evaluated at the argument \(x^4 + 2\) times \(x^3\), which aside from a factor of 4, is the derivative of the argument \(x^4 + 2\).
- Hence we set \(u = x^4 + 2\) and then substitute \(dx \rightarrow \frac{1}{4u(x)} \, du = \frac{1}{4x^3} \, du\).
- Before proceeding further, we should note that this is an indefinite integral so we don’t have to worry about the limits of integration. However we do need to make sure our answer is a function of \(x\) — we cannot leave it as a function of \(u\).
- With this choice of \(u\), the integral then becomes
  \[
  \int x^3 \cos (x^4 + 2) \, dx = \int x^3 \cos(u) \frac{1}{4x^3} \, du \bigg|_{u=x^4+2} = \frac{1}{4} \cos(u) \bigg|_{u=x^4+2} = \frac{1}{4} \sin(u) + C \bigg|_{u=x^4+2} = \frac{1}{4} \sin(x^4 + 2) + C.
  \]

Example 1.4.15 

The next two examples are more involved and require more careful thinking.

Example 1.4.16 \( \int \sqrt{1 + x^2} \, x^3 \, dx \)

Compute \( \int \sqrt{1 + x^2} \, x^3 \, dx \).

- An obvious choice of \(u\) is the argument inside the square root. So substitute \(u = 1 + x^2\) and \(dx \rightarrow \frac{1}{2x} \, du\).
- When we do this we obtain
  \[
  \int \sqrt{1 + x^2} \cdot x^3 \, dx = \int \sqrt{u} \cdot x^3 \cdot \frac{1}{2x} \, du = \int \frac{1}{2} \sqrt{u} \cdot x^2 \, du.
  \]
Unlike all our previous examples, we have not cancelled out all of the $x$’s from the integrand. However before we do the integral with respect to $u$, the integrand must be expressed solely in terms of $u$ — no $x$’s are allowed. (Look that integrand on the right hand side of Theorem 1.4.2.)

- But all is not lost. We can rewrite the factor $x^2$ in terms of the variable $u$. We know that $u = 1 + x^2$, so this means $x^2 = u - 1$. Substituting this into our integral gives

$$\int \sqrt{1 + x^2} \cdot x^3 \, dx = \int \frac{1}{2} \sqrt{u} \cdot x^2 \, du$$

$$= \int \frac{1}{2} \sqrt{u} \cdot (u - 1) \, du$$

$$= \frac{1}{2} \int \left( u^{3/2} - u^{1/2} \right) \, du$$

$$= \frac{1}{2} \left( \frac{2}{5}u^{5/2} - \frac{2}{3}u^{3/2} \right) \bigg|_{u=x^2+1} + C$$

$$= \left( \frac{1}{5}u^{5/2} - \frac{1}{3}u^{3/2} \right) \bigg|_{u=x^2+1} + C$$

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

Oof!

- Don’t forget that you can always check the answer by differentiating:

$$\frac{d}{dx} \left( \frac{1}{5}(x^2+1)^{5/2} - \frac{1}{3}(x^2+1)^{3/2} + C \right) = \frac{d}{dx} \left( \frac{1}{5}(x^2+1)^{5/2} \right) - \frac{d}{dx} \left( \frac{1}{3}(x^2+1)^{3/2} \right)$$

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

$$= x(x^2+1)^{3/2} - x(x^2+1)^{1/2}$$

$$= x \left[ (x^2+1)^{1/2} - (x^2+1)^{-1/2} \right] \cdot \sqrt{x^2+1}$$

$$= x^3 \sqrt{x^2+1}.$$  

which is the original integrand $\checkmark$.

---

**Example 1.4.16**

**Example 1.4.17** ($\int \tan x \, dx$)

Evaluate the indefinite integral $\int \tan(x) \, dx$.

**Solution.**

- At first glance there is nothing to manipulate here and so very little to go on. However we can rewrite $\tan x$ as $\frac{\sin x}{\cos x}$, making the integral $\int \frac{\sin x}{\cos x} \, dx$. This gives us more to work with.
• Now think of the integrand as being the product $\frac{1}{\cos x} \cdot \sin x$. This suggests that we set $u = \cos x$ and that we interpret the first factor as the function “one over” evaluated at $u = \cos x$.

• Substitute $u = \cos x$ and $dx \rightarrow \frac{1}{-\sin x} du$ to give:

$$\int \frac{\sin x}{\cos x} dx = \int \left( \frac{\sin x}{u} - \sin x \right) \frac{1}{u} du \bigg|_{u=\cos x}$$

$$= \int \left( -\frac{1}{u} du \right) \bigg|_{u=\cos x}$$

$$= -\log |\cos x| + C \quad \text{and if we want to go further}$$

$$= \log \left| \frac{1}{\cos x} \right| + C$$

$$= \log |\sec x| + C.$$  

---

**Example 1.4.17**

---

### 1.5 Area between curves

Before we continue our exploration of different methods for integrating functions, we have now have sufficient tools to examine some simple applications of definite integrals. One of the motivations for our definition of “integral” was the problem of finding the area between some curve and the $x$–axis for $x$ running between two specified values. More precisely

$$\int_{a}^{b} f(x) dx$$

is equal to the signed area between the the curve $y = f(x)$, the $x$-axis, and the vertical lines $x = a$ and $x = b$.

We found the area of this region by approximating it by the union of tall thin rectangles, and then found the exact area by taking the limit as the width of the approximating rectangles went to zero. We can use the same strategy to find areas of more complicated regions in the $xy$-plane.

As a preview of the material to come, let $f(x) > g(x) > 0$ and $a < b$ and suppose that we are interested in the area of the region

$$S_1 = \{ (x,y) \mid a \leq x \leq b, g(x) \leq y \leq f(x) \}$$

that is sketched in the left hand figure below.
We already know that \( \int_a^b f(x) \, dx \) is the area of the region
\[ S_2 = \{ (x, y) \mid a \leq x \leq b, 0 \leq y \leq f(x) \} \]
sketched in the middle figure above and that \( \int_a^b g(x) \, dx \) is the area of the region
\[ S_3 = \{ (x, y) \mid a \leq x \leq b, 0 \leq y \leq g(x) \} \]
sketched in the right hand figure above. Now the region \( S_1 \) of the left hand figure can be constructed by taking the region \( S_2 \) of center figure and removing from it the region \( S_3 \) of the right hand figure. So the area of \( S_1 \) is exactly
\[
\int_a^b f(x) \, dx - \int_a^b g(x) \, dx = \int_a^b (f(x) - g(x)) \, dx
\]
This computation depended on the assumption that \( f(x) > g(x) \) and, in particular, that the curves \( y = g(x) \) and \( y = f(x) \) did not cross. If they do cross, as in this figure

then we have to be a lot more careful. The idea is to separate the domain of integration depending on where \( f(x) - g(x) \) changes sign — i.e. where the curves intersect. We will illustrate this in Example 1.5.5 below.

Let us start with an example that makes the link to Riemann sums and definite integrals quite explicit.

**Example 1.5.1**

Find the area bounded by the curves \( y = 4 - x^2, y = x, x = -1 \) and \( x = 1 \).

*Solution.*
Before we do any calculus, it is a very good idea to make a sketch of the area in question. The curves \( y = x, \ x = -1 \) and \( x = 1 \) are all straight lines, while the curve \( y = 4 - x^2 \) is a parabola whose apex is at \((0, 4)\) and then curves down (because of the minus sign in \(-x^2\)) with \(x\)-intercepts at \((\pm 2, 0)\). Putting these together gives

\[
\begin{align*}
\text{We are to find the area of the shaded region. Each point } (x, y) \text{ in this shaded region has } -1 \leq x \leq 1 \text{ and } x \leq y \leq 4 - x^2. \text{ When we were defining the integral (way back in Definition 1.1.9) we used } a \text{ and } b \text{ to denote the smallest and largest allowed values of } x; \text{ let’s do that here too. Let’s also use } B(x) \text{ to denote the bottom curve (i.e. to denote the smallest allowed value of } y \text{ for a given } x) \text{ and use } T(x) \text{ to denote the top curve (i.e. to denote the largest allowed value of } y \text{ for a given } x). \text{ So in this example}
\end{align*}
\]

\[
\begin{align*}
a & = -1 \\
b & = 1 \\
B(x) & = x \\
T(x) & = 4 - x^2
\end{align*}
\]

and the shaded region is

\[
\{ (x, y) \mid a \leq x \leq b, B(x) \leq y \leq T(x) \}
\]

- We use the same strategy as we used when defining the integral in Section 1.1.3:
  - Pick a natural number \( n \) (that we will later send to infinity), then
  - subdivide the region into \( n \) narrow slices, each of width \( \Delta x = \frac{b-a}{n} \).
  - For each \( i = 1, 2, \ldots, n \), slice number \( i \) runs from \( x = x_{i-1} \) to \( x = x_i \), and we approximate its area by the area of a rectangle. We pick a number \( x^*_i \) between \( x_{i-1} \) and \( x_i \) and approximate the slice by a rectangle whose top is at \( y = T(x^*_i) \) and whose bottom is at \( y = B(x^*_i) \).
  - Thus the area of slice \( i \) is approximately \([T(x^*_i) - B(x^*_i)]\Delta x\) (as shown in the figure below).
• So the Riemann sum approximation of the area is

\[ \text{Area} \approx \sum_{i=1}^{n} \left[ T(x_i^*) - B(x_i^*) \right] \Delta x \]

• By taking the limit as \( n \to \infty \) (i.e. taking the limit as the width of the rectangles goes to zero), we convert the Riemann sum into a definite integral (see Definition 1.1.9) and at the same time our approximation of the area becomes the exact area:

\[
\lim_{n \to \infty} \sum_{i=1}^{n} \left[ T(x_i^*) - B(x_i^*) \right] \Delta x = \int_{a}^{b} \left[ T(x) - B(x) \right] \, dx
\]

\[
= \int_{-1}^{1} \left[ 4 - x^2 - x \right] \, dx
\]

\[
= \int_{-1}^{1} \left[ 4 - x - x^2 \right] \, dx
\]

\[
= \left[ 4x - \frac{x^2}{2} - \frac{x^3}{3} \right]_{-1}^{1}
\]

\[
= \left( 4 - \frac{1}{2} - \frac{1}{3} \right) - \left( -4 - \frac{1}{2} + \frac{1}{3} \right)
\]

\[
= \frac{24 - 3 - 2}{6} - \frac{-24 - 3 + 2}{6}
\]

\[
= \frac{19}{6} + \frac{25}{6}
\]

\[
= \frac{44}{6} = \frac{22}{3}.
\]

Oof! Thankfully we generally do not need to go through the Riemann sum steps to get to the answer. Usually, provided we are careful to check where curves intersect and
which curve lies above which, we can just jump straight to the integral

\[
    \text{Area} = \int_a^b [T(x) - B(x)] \, dx. \tag{1.5.1}
\]

So let us redo the above example.

**Example 1.5.2 (Example 1.5.1 revisited)**

Find the area bounded by the curves \( y = 4 - x^2, y = x, x = -1 \) and \( x = 1 \).

**Solution.**

- We first sketch the region

![Image of the region]

and verify\(^{36}\) that \( y = T(x) = 4 - x^2 \) lies above the curve \( y = B(x) = x \) on the region \(-1 \leq x \leq 1\).

- The area between the curves is then

\[
    \text{Area} = \int_a^b [T(x) - B(x)] \, dx \\
    = \int_{-1}^1 [4 - x - x^2] \, dx \\
    = \left[ 4x - \frac{x^2}{2} - \frac{x^3}{3} \right]_{-1}^1 \\
    = \frac{19}{6} + \frac{25}{6} = \frac{44}{6} = \frac{22}{3}.
\]

**Example 1.5.3**

Find the area of the finite region bounded by \( y = x^2 \) and \( y = 6x - 2x^2 \).

\(^{36}\) We should do this by checking where the curves intersect; that is by solving \( T(x) = B(x) \) and seeing if any of the solutions lie in the range \(-1 \leq x \leq 1\).
Solution. This is a little different from the previous question, since we are not given bounding lines \( x = a \) and \( x = b \) — instead we have to determine the minimum and maximum allowed values of \( x \) by determining where the curves intersect. Hence our very first task is to get a good idea of what the region looks like by sketching it.

- Start by sketching the region:
  - The curve \( y = x^2 \) is a parabola. The point on this parabola with the smallest \( y \)-coordinate is \((0, 0)\). As \( |x| \) increases, \( y \) increases so the parabola opens upward.
  - The curve \( y = 6x - 2x^2 = -2(x^2 - 3x) = -2(x - \frac{3}{2})^2 + \frac{9}{2} \) is also a parabola. The point on this parabola with the largest value of \( y \) has \( x = \frac{3}{2} \) (so that the negative term in \(-2(x - \frac{3}{2})^2 + \frac{9}{2}\) is zero). So the point with the largest value of \( y \) is \((\frac{3}{2}, \frac{9}{2})\). As \( x \) moves away from \( \frac{3}{2} \), either to the right or to the left, \( y \) decreases. So the parabola opens downward. The parabola crosses the \( x \)-axis when \( 0 = 6x - 2x^2 = 2x(3 - x) \). That is, when \( x = 0 \) and \( x = 3 \).
  - The two parabolas intersect when \( x^2 = 6x - 2x^2 \), or
    \[
    3x^2 - 6x = 0 \\
    3x(x - 2) = 0
    \]
    So there are two points of intersection, one being \( x = 0, y = 0^2 = 0 \) and the other being \( x = 2, y = 2^2 = 4 \).
  - The finite region between the curves lies between these two points of intersection.

This leads us to the sketch

![Sketch of the region between the curves](image)

- So on this region we have \( 0 \leq x \leq 2 \), the top curve is \( T(x) = 6x - x^2 \) and the bottom
INTEGRATION

1.5 Area between curves

curve is \( B(x) = x^2 \). Hence the area is given by

\[
\text{Area} = \int_a^b [T(x) - B(x)] \, dx
\]

\[
= \int_0^2 [(6x - 2x^2) - (x^2)] \, dx
\]

\[
= \int_0^2 [6x - 3x^2] \, dx
\]

\[
= \left[ 6\frac{x^2}{2} - 3\frac{x^3}{3} \right]_0
\]

\[
= 3(2)^2 - 2^3 = 4
\]

Example 1.5.3

Example 1.5.4

Find the area of the finite region bounded by \( y^2 = 2x + 6 \) and \( y = x - 1 \).

Solution. We show two different solutions to this problem. The first takes the approach we have in Example 1.5.3 but leads to messy algebra. The second requires a little bit of thinking at the beginning but then is quite straightforward. Before we get to that we should start by by sketching the region.

- The curve \( y^2 = 2x + 6 \), or equivalently \( x = \frac{1}{2}y^2 - 3 \) is a parabola. The point on this parabola with the smallest \( x \)-coordinate has \( y = 0 \) (so that the positive term in \( \frac{1}{2}y^2 - 3 \) is zero). So the point on this parabola with the smallest \( x \)-coordinate is \((-3, 0)\). As \(|y|\) increases, \( x \) increases so the parabola opens to the right.

- The curve \( y = x - 1 \) is a straight line of slope 1 that passes through \( x = 1, y = 0 \).

- The two curves intersect when \( \frac{y^2}{2} - 3 = y + 1 \), or

\[
\begin{align*}
y^2 - 6 &= 2y + 2 \\
y^2 - 2y - 8 &= 0 \\
(y + 2)(y - 4) &= 0
\end{align*}
\]

So there are two points of intersection, one being \( y = 4, x = 4 + 1 = 5 \) and the other being \( y = -2, x = -2 + 1 = -1 \).

- Putting this all together gives us the sketch
As noted above, we can find the area of this region by approximating it by a union of narrow vertical rectangles, as we did in Example 1.5.3 — though it is a little harder. The easy way is to approximate it by a union of narrow horizontal rectangles. Just for practice, here is the hard solution. The easy solution is after it.

**Harder solution:**

- As we have done previously, we approximate the region by a union of narrow vertical rectangles, each of width \( \Delta x \). Two of those rectangles are illustrated in the sketch.

- In this region, \( x \) runs from \( a = -3 \) to \( b = 5 \). The curve at the top of the region is

\[
y = T(x) = \sqrt{2x + 6}
\]

The curve at the bottom of the region is more complicated. To the left of \((-1, -2)\) the lower half of the parabola gives the bottom of the region while to the right of \((-1, -2)\) the straight line gives the bottom of the region. So

\[
B(x) = \begin{cases} 
-\sqrt{2x + 6} & \text{if } -3 \leq x \leq -1 \\
x - 1 & \text{if } -1 \leq x \leq 5 
\end{cases}
\]

- Just as before, the area is still given by the formula \( \int_{a}^{b} [T(x) - B(x)] \, dx \), but to accommodate our \( B(x) \), we have to split up the domain of integration when we evaluate.
the integral.

\[
\int_a^b [T(x) - B(x)] \, dx = \int_{-3}^{-1} [T(x) - B(x)] \, dx + \int_{-1}^{5} [T(x) - B(x)] \, dx
\]

\[
= \int_{-3}^{-1} [\sqrt{2x + 6} - (-\sqrt{2x + 6})] \, dx + \int_{-1}^{5} [\sqrt{2x + 6} - (x - 1)] \, dx
\]

\[
= 2 \int_{-3}^{-1} \sqrt{2x + 6} \, dx + \int_{-1}^{5} \sqrt{2x + 6} - \int_{-1}^{5} (x - 1) \, dx
\]

- The third integral is straightforward, while we evaluate the first two via the substitution rule. In particular, set \( u = 2x + 6 \) and replace \( dx \to \frac{1}{2}du \). Also \( u(-3) = 0, u(-1) = 4, u(5) = 16 \). Hence

\[
\text{Area} = 2 \int_0^4 \sqrt{u} \, \frac{du}{2} + \int_4^{16} \sqrt{u} \, \frac{du}{2} - \int_{-1}^{5} (x - 1) \, dx
\]

\[
= 2 \left[ \frac{u^{3/2}}{3/2} \right]_0^4 + \left[ \frac{u^{3/2}}{3/2} \right]_4^{16} - \left[ \frac{x^2}{2} - x \right]_{-1}^{5}
\]

\[
= \frac{2}{3} [8 - 0] + \frac{1}{3} [64 - 8] - \left[ \left( \frac{25}{2} - 5 \right) - \left( \frac{1}{2} + 1 \right) \right]
\]

\[
= \frac{72}{3} - \frac{24}{2} + 6 = 18
\]

**Easier solution:**
The easy way to determine the area of our region is to approximate by narrow horizontal rectangles, rather than narrow vertical rectangles. (Really we are just swapping the roles of \( x \) and \( y \) in this problem)

- Look at our sketch of the region again — each point \((x, y)\) in our region has \(-2 \leq y \leq 4\) and \( \frac{1}{2}(y^2 - 6) \leq x \leq y + 1 \).

- Let’s use
  - \( c \) to denote the smallest allowed value of \( y \),
  - \( d \) to denote the largest allowed value of \( y \)
  - \( L(y) \) (“\( L \)” stands for “left”) to denote the smallest allowed value of \( x \), when the \( y \)-coordinate is \( y \), and
  - \( R(y) \) (“\( R \)” stands for “right”) to denote the largest allowed value of \( x \), when the \( y \)-coordinate is \( y \).

So, in this example,

\[
\begin{align*}
c &= -2 & d &= 4 & L(y) &= \frac{1}{2}(y^2 - 6) & R(y) &= y + 1
\end{align*}
\]
and the shaded region is
\[ \{ (x, y) \mid c \leq y \leq d, L(y) \leq x \leq R(y) \} \]

- Our strategy is now nearly the same as that used in Example 1.5.1:
  - Pick a natural number \( n \) (that we will later send to infinity), then
  - subdivide the interval \( c \leq y \leq d \) into \( n \) narrow subintervals, each of width \( \Delta y = \frac{d-c}{n} \). Each subinterval cuts a thin horizontal slice from the region (see the figure below).
  - We approximate the area of slice number \( i \) by the area of a thin horizontal rectangle (indicated by the dark rectangle in the figure below). On this slice, the \( y \)-coordinate runs over a very narrow range. We pick a number \( y_i^* \), somewhere in that range. We approximate slice \( i \) by a rectangle whose left side is at \( x = L(y_i^*) \) and whose right side is at \( x = R(y_i^*) \).
  - Thus the area of slice \( i \) is approximately \( [R(x_i^*) - L(x_i^*)] \Delta y \).

\[
\begin{align*}
\text{The desired area is} \quad & \quad \lim_{n \to \infty} \sum_{i=1}^{n} [R(y_i^*) - L(y_i^*)] \Delta y = \int_{c}^{d} [R(y) - L(y)] \, dy \\
& \quad \text{Riemann sum \to integral} \\
& \quad = \int_{-2}^{4} [y + 1 - \frac{1}{2}(y^2 - 6)] \, dy \\
& \quad = \int_{-2}^{4} [-\frac{1}{2}y^2 + y + 4] \, dy \\
& \quad = \left[ -\frac{1}{6}y^3 + \frac{1}{2}y^2 + 4y \right]_{-2}^{4} \\
& \quad = -\frac{1}{6}(64 - (-8)) + \frac{1}{2}(16 - 4) + 4(4 + 2) \\
& \quad = -12 + 6 + 24 \\
& \quad = 18
\end{align*}
\]
One last example.

**Example 1.5.5**

Find the area between the curves \( y = \frac{1}{\sqrt{2}} \) and \( y = \sin(x) \) with \( x \) running from 0 to \( \frac{\pi}{2} \).

*Solution.* This one is a little trickier since (as we shall see) the region is split into two pieces and we need to treat them separately.

- Again we start by sketching the region.

We want the shaded area.

- Unlike our previous examples, the bounding curves \( y = \frac{1}{\sqrt{2}} \) and \( y = \sin(x) \) cross in the middle of the region of interest. They cross when \( y = \frac{1}{\sqrt{2}} \) and \( \sin(x) = y = \frac{1}{\sqrt{2}} \), i.e. when \( x = \frac{\pi}{4} \). So

  - to the left of \( x = \frac{\pi}{4} \), the top boundary is part of the straight line \( y = \frac{1}{\sqrt{2}} \) and the bottom boundary is part of the curve \( y = \sin(x) \)

  - while to the right of \( x = \frac{\pi}{4} \), the top boundary is part of the curve \( y = \sin(x) \) and the bottom boundary is part of the straight line \( y = \frac{1}{\sqrt{2}} \).

- Thus the formulae for the top and bottom boundaries are

\[
T(x) = \begin{cases} 
\frac{1}{\sqrt{2}} & \text{if } 0 \leq x \leq \frac{\pi}{4} \\
\sin(x) & \text{if } \frac{\pi}{4} \leq x \leq \frac{\pi}{2} 
\end{cases} 
\quad B(x) = \begin{cases} 
\sin(x) & \text{if } 0 \leq x \leq \frac{\pi}{4} \\
\frac{1}{\sqrt{2}} & \text{if } \frac{\pi}{4} \leq x \leq \frac{\pi}{2} 
\end{cases}
\]

We may compute the area of interest using our canned formula

\[
\text{Area} = \int_{a}^{b} \left[ T(x) - B(x) \right] \, dx
\]
but since the formulas for \( T(x) \) and \( B(x) \) change at the point \( x = \pi/4 \), we must split the domain of the integral in two at that point:  

- Our integral over the domain \( 0 \leq x \leq \pi/2 \) is split into an integral over \( 0 \leq x \leq \pi/4 \) and one over \( \pi/4 \leq x \leq \pi/2 \):

\[
\text{Area} = \int_0^{\pi/2} [T(x) - B(x)] \, dx \\
= \int_0^{\pi/4} [T(x) - B(x)] \, dx + \int_{\pi/4}^{\pi/2} [T(x) - B(x)] \, dx \\
= \int_0^{\pi/4} \left[ \frac{1}{\sqrt{2}} - \sin(x) \right] \, dx + \int_{\pi/4}^{\pi/2} \sin(x) - \frac{1}{\sqrt{2}} \right] \, dx \\
= \left[ \frac{x}{\sqrt{2}} + \cos(x) \right]_0^{\pi/4} + \left[ -\cos(x) - \frac{x}{\sqrt{2}} \right]_{\pi/4}^{\pi/2} \\
= \left[ \frac{1}{\sqrt{2}} \left( \frac{\pi}{4} \right) - \cos \left( \frac{\pi}{4} \right) - 1 \right] + \left[ \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} \left( \frac{\pi}{4} \right) \right] \\
= \frac{2}{\sqrt{2}} - 1 \\
= \sqrt{2} - 1
\]

### 1.6 Volumes

Another simple application of integration is computing volumes. We use the same strategy as we used to express areas of regions in two dimensions as integrals — approximate the region by a union of small, simple pieces whose volume we can compute and then take the limit as the “piece size” tends to zero.

In many cases this will lead to “multivariable integrals” that are beyond our present scope. But there are some special cases in which this leads to integrals that we can handle. Here are some examples.

**Example 1.6.1 (Cone)**

Find the volume of the circular cone of height \( h \) and radius \( r \).

---

37 We are effectively computing the area of the region by computing the area of the two disjoint pieces separately. Alternatively, if we set \( f(x) = \sin(x) \) and \( g(x) = \frac{1}{\sqrt{2}} \), we can rewrite the integral \( \int_0^\pi [T(x) - B(x)] \, dx \) as \( \int_0^\pi |f(x) - g(x)| \, dx \). To see that the two integrals are the same, split the domain of integration where \( f(x) - g(x) \) changes sign.

38 Well — arguably the idea isn’t too complicated and is a continuation of the idea used to compute areas in the previous section. In practice this can be quite tricky as we shall see.

39 Typically such integrals (and more) are covered in a third calculus course.
Solution. Here is a sketch of the cone. We have called the vertical axis $x$, just so that we end up with a “$dx$” integral.

- In what follows we will slice the cone into thin horizontal “pancakes”. In order to approximate the volume of those slices, we need to know the radius of the cone at a height $x$ above its point. Consider the cross sections shown in the following figure.

At full height $h$, the cone has radius $r$. If we cut the cone at height $x$, then by similar triangles (see the figure on the right) the radius will be $\frac{x}{h} \cdot r$.

- Now think of cutting the cone into $n$ thin horizontal “pancakes”. Each such pancake is approximately a squat cylinder of height $\Delta x = \frac{h}{n}$. This is very similar to how we approximated the area under a curve by $n$ tall thin rectangles. Just as we approximated the area under the curve by summing these rectangles, we can approximate the volume of the cone by summing the volumes of these cylinders. Here is a side view of the cone and one of the cylinders.
• We follow the method we used in Example 1.5.1, except that our slices are now pancakes instead of rectangles.

  – Pick a natural number $n$ (that we will later send to infinity), then
  – subdivide the cone into $n$ thin pancakes, each of width $\Delta x = \frac{h}{n}$.
  – For each $i = 1, 2, \ldots, n$, pancake number $i$ runs from $x = x_{i-1} = (i - 1) \cdot \Delta x$ to $x = x_i = i \cdot \Delta x$, and we approximate its volume by the volume of a squat cone. We pick a number $x^*_i$ between $x_{i-1}$ and $x_i$ and approximate the pancake by a cylinder of height $\Delta x$ and radius $\frac{x^*_i}{h} r$.

  – Thus the volume of pancake $i$ is approximately $\pi \left( \frac{x^*_i}{h} r \right)^2 \Delta x$ (as shown in the figure above).

• So the Riemann sum approximation of the volume is

$$\text{Area} \approx \sum_{i=1}^{n} \pi \left( \frac{x^*_i}{h} r \right)^2 \Delta x$$

• By taking the limit as $n \to \infty$ (i.e. taking the limit as the thickness of the pancakes goes to zero), we convert the Riemann sum into a definite integral (see Definition 1.1.9) and at the same time our approximation of the volume becomes the exact volume:

$$\int_{0}^{h} \pi \left( \frac{x}{h} r \right)^2 \, dx$$

Our life\textsuperscript{40} would be easier if we could avoid all this formal work with Riemann sums every time we encounter a new volume. So before we compute the above integral, let us redo the above calculation in a less formal manner.

• Start again from the picture of the cone and think of slicing it into thin pancakes,

\[\text{each of width } dx.\]

\textsuperscript{40} At least the bits of it involving integrals.
The pancake at height \( x \) above the point of the cone (which is the fraction \( \frac{x}{h} \) of the total height of the cone) has

- radius \( \frac{x}{h} \cdot r \) (the fraction \( \frac{x}{h} \) of the full radius, \( r \)) and so
- cross-sectional area \( \pi \left( \frac{x}{h} r \right)^2 \),
- thickness \( dx \) — we have done something a little sneaky here, see the discussion below.
- volume \( \pi \left( \frac{x}{h} r \right)^2 dx \)

As \( x \) runs from 0 to \( h \), the total volume is

\[
\int_0^h \pi \left( \frac{x}{h} r \right)^2 \, dx = \frac{\pi r^2}{h^2} \int_0^h x^2 \, dx
\]

\[
= \frac{\pi r^2}{h^2} \left[ \frac{x^3}{3} \right]_0^h
\]

\[
= \frac{1}{3} \pi r^2 h
\]

In this second computation we are using a time-saving trick. As we saw in the formal computation above, what we really need to do is pick a natural number \( n \), slice the cone into \( n \) pancakes each of thickness \( \Delta x = \frac{h}{n} \) and then take the limit as \( n \to \infty \). This led to the Riemann sum

\[
\sum_{i=1}^n \pi \left( \frac{x_i^*}{h} r \right)^2 \Delta x
\]

which becomes \( \int_0^h \pi \left( \frac{x}{h} r \right)^2 \, dx \)

So knowing that we will replace

\[
\sum_{i=1}^n \rightarrow \int_0^h
\]

\[
x_i^* \rightarrow x
\]

\[
\Delta x \rightarrow dx
\]

when we take the limit, we have just skipped the intermediate steps. While this is not entirely rigorous, it can be made so, and does save us a lot of algebra.
Example 1.6.2 (Sphere)

Find the volume of the sphere of radius $r$.

**Solution.** We’ll find the volume of the part of the sphere in the first octant, sketched below. Then we’ll multiply by 8.

- To compute the volume, we slice it up into thin vertical “pancakes” (just as we did in the previous example).
- Each pancake is one quarter of a thin circular disk. The pancake a distance $x$ from the $yz$–plane is shown in the sketch above. The radius of that pancake is the distance from the dot shown in the figure to the $x$–axis, i.e. the $y$–coordinate of the dot. To get the coordinates of the dot, observe that
  - it lies the $xy$–plane, and so has $z$–coordinate zero, and that
  - it also lies on the sphere, so that its coordinates obey $x^2 + y^2 + z^2 = r^2$. Since $z = 0$ and $y > 0$, $y = \sqrt{r^2 - x^2}$.
- So the pancake at distance $x$ from the $yz$–plane has
  - thickness $dx$ and
  - radius $\sqrt{r^2 - x^2}$
  - cross–sectional area $\frac{1}{4} \pi (\sqrt{r^2 - x^2})^2$ and hence
  - volume $\frac{\pi}{4} (r^2 - x^2) dx$

---

41 The first octant is the set of all points $(x, y, z)$ with $x \geq 0$, $y \geq 0$ and $z \geq 0$.
42 Yet again what we really do is pick a natural number $n$, slice the octant of the sphere into $n$ pancakes each of thickness $\Delta x = \frac{r}{n}$ and then take the limit $n \to \infty$. In the integral $\Delta x$ is replaced by $dx$. Knowing that this is what is going to happen, we again just skip a few steps.
• As $x$ runs from 0 to $r$, the total volume of the part of the sphere in the first octant is

$$\int_0^r \frac{\pi}{4} (r^2 - x^2) \, dx = \frac{\pi}{4} \left[ r^2 x - \frac{x^3}{3} \right]_0^r = \frac{1}{6} \pi r^3$$

and the total volume of the whole sphere is eight times that, which is $\frac{4}{3} \pi r^3$, as expected.

Example 1.6.2

Example 1.6.3 (Revolving a region)

The region between the lines $y = 3$, $y = 5$, $x = 0$ and $x = 4$ is rotated around the line $y = 2$. Find the volume of the region swept out.

Solution. As with most of these problems, we should start by sketching the problem.

Consider the region and slice it into thin vertical strips of width $dx$.

Now we are to rotate this region about the line $y = 2$. Imagine looking straight down the axis of rotation, $y = 2$, end on. The symbol in the figure above just to the right of the end the line $y = 2$ is supposed to represent your eye. Here is what you see as the rotation takes place.

43 Okay okay… We missed the pupil. I’m sure there is a pun in there somewhere.
Upon rotation about the line $y = 2$ our strip sweeps out a “washer”
- whose cross-section is a disk of radius $5 - 2 = 3$ from which a disk of radius $3 - 2 = 1$ has been removed so that it has a
- cross-sectional area of $\pi 3^2 - \pi 1^2 = 8\pi$ and a
- thickness $dx$ and hence a
- volume $8\pi \, dx$.

As our leftmost strip is at $x = 0$ and our rightmost strip is at $x = 4$, the total
$$\text{Volume} = \int_0^4 8\pi \, dx = (8\pi)(4) = 32\pi$$

Notice that we could also reach this answer by writing the volume as the difference of two cylinders.
- The outer cylinder has radius $(5 - 2)$ and length $4$. This has volume
  $$V_{outer} = \pi r^2 \ell = \pi \cdot 3^2 \cdot 4 = 36\pi.$$  
- The inner cylinder has radius $(3 - 2)$ and length $4$. This has volume
  $$V_{inner} = \pi r^2 \ell = \pi \cdot 1^2 \cdot 4 = 4\pi.$$  
- The volume we want is the difference of these two, namely
  $$V = V_{outer} - V_{inner} = 32\pi.$$  

Let us turn up the difficulty a little on this last example.

The region between the curve $y = \sqrt{x}$, and the lines $y = 0$, $x = 0$ and $x = 4$ is rotated around the line $y = 0$. Find the volume of the region swept out.

Solution. We can approach this in much the same way as the previous example.
- Consider the region and cut it into thin vertical strips of width $dx$. 

![Diagram of revolving region]
• When we rotate the region about the line $y = 0$, each strip sweeps out a thin pancake
  - whose cross-section is a disk of radius $\sqrt{x}$ with a
  - cross-sectional area of $\pi(\sqrt{x})^2 = \pi x$ and a
  - thickness $dx$ and hence a
  - volume $\pi x dx$.

• As our leftmost strip is at $x = 0$ and our rightmost strip is at $x = 4$, the total
  
  $$
  \text{Volume} = \int_0^4 \pi x \, dx = \left. \left[ \frac{\pi}{2} x^2 \right] \right|_0^4 = 8\pi
  $$

In the last example we considered rotating a region around the $x$-axis. Let us do the same but rotating around the $y$-axis.

The region between the curve $y = \sqrt{x}$, and the lines $y = 0$, $x = 0$ and $x = 4$ is rotated around the line $x = 0$. Find the volume of the region swept out.

Solution.

• We will cut the region into horizontal slices, so we should write $x$ as a function of $y$.
  That is, the region is bounded by $x = y^2$, $x = 4$, $y = 0$ and $y = 2$.

• Now slice the region into thin horizontal strips of width $dy$.

• When we rotate the region about the line $y = 0$, each strip sweeps out a thin washer
  - whose inner radius is $y^2$ and outer radius is 4, and
  - thickness is $dy$ and hence
  - has volume $\pi(r_{out}^2 - r_{in}^2)dy = \pi(16 - y^4)dy$. 
As our bottommost strip is at \( y = 0 \) and our topmost strip is at \( y = 2 \), the total volume

\[
\text{Volume} = \int_0^2 \pi (16 - y^4) \, dy = \left[ 16\pi y - \frac{\pi y^5}{5} \right]_0^2 = 32\pi - \frac{32\pi}{5} = \frac{128\pi}{5}.
\]

There is another way\(^4^4\) to do this one which we show at the end of this section.

**Example 1.6.6 (Pyramid)**

Find the volume of the pyramid which has height \( h \) and whose base is a square of side \( b \).

*Solution.* Here is a sketch of the part of the pyramid that is in the first octant; we display only this portion to make the diagrams simpler. Note that this diagram shows only 1 quarter of the whole pyramid.

- To compute its volume, we slice it up into thin horizontal “square pancakes”. A typical pancake also appears in the sketch above.
  - The pancake at height \( z \) is the fraction \( \frac{h-z}{h} \) of the distance from the peak of the pyramid to its base.
  - So the full pancake\(^4^5\) at height \( z \) is a square of side \( \frac{h-z}{h} b \). As a check, note that when \( z = h \) the pancake has side \( \frac{h-h}{h} b = 0 \), and when \( z = 0 \) the pancake has side \( \frac{h-0}{h} b = b \).
  - So the pancake has cross-sectional area \( \left( \frac{h-z}{h} b \right)^2 \) and thickness\(^4^6\) \( dz \) and hence
    - volume \( \left( \frac{h-z}{h} b \right)^2 \, dz \).

\(^{4^4}\) The method is not a core part of the course and should be considered optional.

\(^{4^5}\) Note that this is the full pancake, not just the part in the first octant.

\(^{4^6}\) We are again using our Riemann sum avoiding trick.
• The volume of the whole pyramid (not just the part of the pyramid in the first octant) is
\[
\int_0^h \left( \frac{h - z}{h} b \right)^2 dz = \frac{b^2}{h^2} \int_0^h (h - z)^2 dz
\]
\[
= \frac{b^2}{h^2} \int_0^h -t^2 dt \quad \text{substitution rule with } t = (h - z), dz \rightarrow -dt
\]
\[
= -\frac{b^2}{h^2} \left[ \frac{t^3}{3} \right]_0^h
\]
\[
= -\frac{b^2}{h^2} \left[ -\frac{h^3}{3} \right]
\]
\[
= \frac{1}{3} b^2 h
\]

Let’s ramp up the difficulty a little.

Example 1.6.7 (Napkin Ring)

Suppose you make two napkin rings by drilling holes with different diameters through two wooden balls. One ball has radius \( r \) and the other radius \( R \) with \( r < R \). You choose the diameter of the holes so that both napkin rings have the same height, \( 2h \). See the figure below.

Which ring has more wood in it?

Solution. We’ll compute the volume of the napkin ring with radius \( R \). We can then obtain the volume of the napkin ring of radius \( r \), by just replacing \( R \rightarrow r \) in the result.

• To compute the volume of the napkin ring of radius \( R \), we slice it up into thin horizontal “pancakes”. Here is a sketch of the part of the napkin ring in the first octant showing a typical pancake.

47 Handy things to have (when combined with cloth napkins) if your parents are coming to dinner and you want to convince them that you are “taking care of yourself”.

48 A good question to ask to distract your parents from the fact you are serving frozen burritos.
• The coordinates of the two points marked in the $yz$–plane of that figure are found by remembering that
  – the equation of the sphere is $x^2 + y^2 + z^2 = R^2$.
  – The two points have $y > 0$ and are in the $yz$–plane, so that $x = 0$ for them. So $y = \sqrt{R^2 - z^2}$.
  – In particular, at the top of the napkin ring $z = h$, so that $y = \sqrt{R^2 - h^2}$.

• The pancake at height $z$, shown in the sketch, is a “washer” — a circular disk with a circular hole cut in its center.
  – The outer radius of the washer is $\sqrt{R^2 - z^2}$ and
  – the inner radius of the washer is $\sqrt{R^2 - h^2}$. So the
  – cross–sectional area of the washer is
    \[ \pi \left( \sqrt{R^2 - z^2} \right)^2 - \pi \left( \sqrt{R^2 - h^2} \right)^2 = \pi (h^2 - z^2) \]

• The pancake at height $z$
  – has thickness $dz$ and
  – cross–sectional area $\pi (h^2 - z^2)$ and hence
  – volume $\pi (h^2 - z^2)dz$.

• Since $z$ runs from $-h$ to $+h$, the total volume of wood in the napkin ring of radius $R$ is
  \[
  \int_{-h}^{h} \pi (h^2 - z^2) dz = \pi \left[ h^2 z - \frac{z^3}{3} \right]_{-h}^{h} \\
  = \pi \left[ \left( h^3 - \frac{h^3}{3} \right) - \left( (-h)^3 - \frac{(-h)^3}{3} \right) \right] \\
  = \pi \left[ \frac{2}{3} h^3 - \frac{2}{3} (-h)^3 \right] \\
  = \frac{4}{3} \pi h^3
  \]
This volume is independent of $R$. Hence the napkin ring of radius $r$ contains precisely the same volume of wood as the napkin ring of radius $R$!

**Example 1.6.7**

A $45^\circ$ notch is cut to the centre of a cylindrical log having radius 20cm. One plane face of the notch is perpendicular to the axis of the log. See the sketch below. What volume of wood was removed?

**Solution.** We show two solutions to this problem which are of comparable difficulty. The difference lies in the shape of the pancakes we use to slice up the volume. In solution 1 we cut rectangular pancakes parallel to the $yz$–plane and in solution 2 we slice triangular pancakes parallel to the $xz$–plane.

**Solution 1:**

- Concentrate on the notch. Rotate it around so that the plane face lies in the $xy$–plane.
- Then slice the notch into vertical rectangles (parallel to the $yz$–plane) as in the figure on the left below.

- The cylindrical log had radius 20cm. So the circular part of the boundary of the base of the notch has equation $x^2 + y^2 = 20^2$. (We’re putting the origin of the $xy$–plane at the centre of the circle.) If our coordinate system is such that $x$ is constant on each slice, then
- the base of the slice is the line segment from \((x, -y, 0)\) to \((x, +y, 0)\) where \(y = \sqrt{20^2 - x^2}\) so that
- the slice has width \(2y = 2\sqrt{20^2 - x^2}\) and
- height \(x\) (since the upper face of the notch is at 45° to the base — see the side view sketched in the figure on the right above).
- So the slice has cross-sectional area \(2x\sqrt{20^2 - x^2}\).

- On the base of the notch \(x\) runs from 0 to 20 so the volume of the notch is

\[
V = \int_0^{20} 2x\sqrt{20^2 - x^2} \, dx
\]

Make the change of variables \(u = 20^2 - x^2\) (don’t forget to change \(dx \to -\frac{1}{2x} \, du\)):

\[
V = \int_0^{20} \sqrt{\frac{u^{3/2}}{3/2}} \, du
= \left[ \frac{u^{3/2}}{3/2} \right]_0^{20^2}
= \frac{2}{3} (20^3 - 0)
= \frac{16,000}{3}
\]

Solution 2:

- Concentrate of the notch. Rotate it around so that its base lies in the \(xy\)-plane with the skinny edge along the \(y\)-axis.
- Slice the notch into triangles parallel to the \(xz\)-plane as in the figure on the left below. In the figure below, the triangle happens to lie in a plane where \(y\) is negative.

- The cylindrical log had radius 20cm. So the circular part of the boundary of the base of the notch has equation \(x^2 + y^2 = 20^2\). Our coordinate system is such that \(y\) is constant on each slice, so that
the base of the triangle is the line segment from \((0, y, 0)\) to \((x, y, 0)\) where \(x = \sqrt{20^2 - y^2}\) so that

- the triangle has base \(x = \sqrt{20^2 - y^2}\) and
- height \(x = \sqrt{20^2 - y^2}\) (since the upper face of the notch is at 45° to the base — see the side view sketched in the figure on the right above).

- So the slice has cross-sectional area \(\frac{1}{2} (\sqrt{20^2 - y^2})^2\).

On the base of the notch \(y\) runs from \(-20\) to \(20\), so the volume of the notch is

\[
V = \frac{1}{2} \int_{-20}^{20} (20^2 - y^2) \, dy
\]

\[
= \int_{0}^{20} (20^2 - y^2) \, dy
\]

\[
= \left[ 20^2 y - \frac{y^3}{3} \right]_{0}^{20}
\]

\[
= \frac{2}{3} 20^3 = \frac{16,000}{3}
\]

**Optional — Cylindrical shells**

Let us return to Example 1.6.5 in which we rotate a region around the \(y\)-axis. Here we show another solution to this problem which is obtained by slicing the region into vertical strips. When rotated about the \(y\)-axis, each such strip sweeps out a thin cylindrical shell. Hence the name of this approach (and this subsection).

The region between the curve \(y = \sqrt{x}\), and the lines \(y = 0, x = 0\) and \(x = 4\) is rotated around the line \(x = 0\). Find the volume of the region swept out.

**Solution.**

- Consider the region and cut it into thin vertical strips of width \(dx\).
• When we rotate the region about the line \( y = 0 \), each strip sweeps out a thin cylindrical shell
  - whose radius is \( x \),
  - height is \( \sqrt{x} \), and
  - thickness is \( dx \) and hence
  - has volume \( 2\pi \times \text{radius} \times \text{height} \times \text{thickness} = 2\pi x^{3/2} dx \).
• As our leftmost strip is at \( x = 0 \) and our rightmost strip is at \( x = 4 \), the total

\[
\text{Volume} = \int_{0}^{4} 2\pi x^{3/2} \, dx = \left[ \frac{4\pi}{5} x^{5/2} \right]_{0}^{4} = \frac{4\pi}{5} \cdot 32 = \frac{128\pi}{5}
\]

which (thankfully) agrees with our previous computation.

---

### 1.7 Integration by parts

The fundamental theorem of calculus tells us that it is very easy to integrate a derivative. In particular, we know that

\[
\int \frac{d}{dx} (F(x)) \, dx = F(x) + C
\]

We can exploit this in order to develop another rule for integration — in particular a rule to help us integrate products of simpler function such as

\[
\int xe^x \, dx
\]

In so doing we will arrive at a method called “integration by parts”.

To do this we start with the product rule and integrate. Recall that the product rule says

\[
\frac{d}{dx} u(x)v(x) = u'(x) v(x) + u(x) v'(x)
\]

Integrating this gives

\[
\int [u'(x) v(x) + u(x) v'(x)] \, dx = \left[ \text{a function whose derivative is } u'v + uv' \right] + C
\]

\[
= u(x)v(x) + C
\]

Now this, by itself, is not terribly useful. In order to apply it we need to have a function whose integrand is a sum of products that is in exactly this form \( u'(x)v(x) + u(x)v'(x) \). This is far too specialised.
However if we tease this apart a little:

\[
\int [u'(x) v(x) + u(x) v'(x)] \, dx = \int u'(x) v(x) \, dx + \int u(x) v'(x) \, dx
\]

Bring one of the integrals to the left-hand side

\[u(x)v(x) - \int u'(x) v(x) \, dx = \int u(x) v'(x) \, dx\]

Swap left and right sides

\[\int u(x) v'(x) \, dx = u(x) v(x) - \int u'(x) v(x) \, dx\]

In this form we take the integral of one product and express it in terms of the integral of a different product. If we express it like that, it doesn’t seem too useful. However, if the second integral is easier, then this process helps us.

Let us do a simple example before explaining this more generally.

---

**Example 1.7.1** \(\int xe^x \, dx\)

Compute the integral \(\int xe^x \, dx\).

**Solution.**

- We start by taking the equation above

\[\int u(x) v'(x) \, dx = u(x) v(x) - \int u'(x) v(x) \, dx\]

- Now set \(u(x) = x\) and \(v'(x) = e^x\). How did we know how to make this choice? We will explain some strategies later. For now, let us just accept this choice and keep going.

- In order to use the formula we need to know \(u'(x)\) and \(v(x)\). In this case it is quite straightforward: \(u'(x) = 1\) and \(v(x) = e^x\).

- Plug everything into the formula:

\[\int xe^x \, dx = xe^x - \int e^x \, dx\]

So our original more difficult integral has been turned into a question of computing an easy one.

\[= xe^x - e^x + C\]

- We can check our answer by differentiating:

\[\frac{d}{dx}(xe^x - e^x + C) = xe^x + 1 \cdot e^x - e^x + 0\]

(by product rule)

\[= xe^x\] as required.
The process we have used in the above example is called “integration by parts”. When our integrand is a product we try to write it as \( u(x)v'(x) \) — we need to choose one factor to be \( u(x) \) and the other to be \( v'(x) \). We then compute \( u'(x) \) and \( v(x) \) and then apply the following theorem:

**Theorem 1.7.2 (Integration by parts).**

Let \( u(x) \) and \( v(x) \) be continuously differentiable. Then

\[
\int u(x) v'(x) \, dx = u(x) v(x) - \int v(x) u'(x) \, dx
\]

If we write \( dv \) for \( v'(x) \, dx \) and \( du \) for \( u'(x) \, dx \) (as the substitution rule suggests), then the formula becomes

\[
\int u \, dv = uv - \int v \, du
\]

The application of this formula is known as integration by parts. The corresponding statement for definite integrals is

\[
\int_{a}^{b} u(x) v'(x) \, dx = u(b)v(b) - u(a)v(a) - \int_{a}^{b} v(x) u'(x) \, dx
\]

Integration by parts is not as easy to apply as the product rule for derivatives. This is because it relies on us

(1) judiciously choosing \( u(x) \) and \( v'(x) \), then

(2) computing \( u'(x) \) and \( v(x) \) — which requires us to antidifferentiate \( v'(x) \), and finally

(3) that the integral \( \int u'(x)v(x) \, dx \) is easier than the integral we started with.

Notice that any antiderivative of \( v'(x) \) will do. All antiderivatives of \( v'(x) \) are of the form \( v(x) + A \) with \( A \) a constant. Putting this into the integration by parts formula gives

\[
\int u(x)v'(x) \, dx = u(x)(v(x) + A) - \int u'(x)(v(x) + A) \, dx \\
= u(x)v(x) + Au(x) - \int u'(x)v(x) \, dx - A \int u'(x) \, dx
\]

\[
= u(x)v(x) - \int u'(x)v(x) \, dx + C
\]

So that constant \( A \) will always cancel out.

In most applications (but not all) our integrand will be a product of two factors so we have two choices for \( u(x) \) and \( v'(x) \). Typically one of these choices will be “good” (in that
it results in a simpler integral) while the other will be “bad” (we cannot antidifferentiate our choice of $v'(x)$ or the resulting integral is harder). Let us illustrate what we mean by returning to our previous example.

**Example 1.7.3 ($\int xe^x \, dx$ — again)**

Our integrand is the product of two factors

\[
x \quad \text{and} \quad e^x
\]

This gives us two obvious choices of $u$ and $v'$:

- $u(x) = x$ \quad $v'(x) = e^x$
- or
- $u(x) = e^x$ \quad $v'(x) = x$

We should explore both choices:

1. If take $u(x) = x$ and $v'(x) = e^x$. We then quickly compute

\[
u'(x) = 1 \quad \text{and} \quad v(x) = e^x
\]

which means we will need to integrate (in the right-hand side of the integration by parts formula)

\[
\int u'(x)v(x) \, dx = \int 1 \cdot e^x \, dx
\]

which looks straightforward. This is a good indication that this is the right choice of $u(x)$ and $v'(x)$.

2. But before we do that, we should also explore the other choice, namely $u(x) = e^x$ and $v'(x) = x$. This implies that

\[
u'(x) = e^x \quad \text{and} \quad v(x) = \frac{1}{2} x^2
\]

which means we need to integrate

\[
\int u'(x)v(x) \, dx = \int \frac{1}{2} x^2 \cdot e^x \, dx.
\]

This is at least as hard as the integral we started with. Hence we should try the first choice.

With our choice made, we integrate by parts to get

\[
\int xe^x \, dx = xe^x - \int e^x \, dx
\]

\[
= xe^x - e^x + C.
\]

The above reasoning is a very typical workflow when using integration by parts.

**Example 1.7.3**

Integration by parts is often used
• to eliminate factors of $x$ from an integrand like $xe^x$ by using that $\frac{d}{dx}x = 1$ and
• to eliminate a log $x$ from an integrand by using that $\frac{d}{dx} \log x = \frac{1}{x}$ and
• to eliminate inverse trig functions, like $\arctan x$, from an integrand by using that, for example, $\frac{d}{dx} \arctan x = \frac{1}{1+x^2}$.

Example 1.7.4 ($\int x \sin x \,dx$)

**Solution.**

• Again we have a product of two factors giving us two possible choices.

(1) If we choose $u(x) = x$ and $v'(x) = \sin x$, then we get

$$u'(x) = 1 \quad \text{and} \quad v(x) = -\cos x$$

which is looking promising.

(2) On the other hand if we choose $u(x) = \sin x$ and $v'(x) = x$, then we have

$$u'(x) = \cos x \quad \text{and} \quad v(x) = \frac{1}{2}x^2$$

which is looking worse — we’d need to integrate $\int \frac{1}{2}x^2 \cos x \,dx$.

• So we stick with the first choice. Plugging $u(x) = x$, $v(x) = -\cos x$ into integration by parts gives us

$$\int x \sin x \,dx = -x \cos x - \int 1 \cdot (-\cos x) \,dx$$

$$= -x \cos x + \sin x + C$$

• Again we can check our answer by differentiating:

$$\frac{d}{dx} (-x \cos x + \sin x + C) = -\cos x + x \sin x + \cos x + 0$$

$$= x \sin x \checkmark$$

Once we have practised this a bit we do not really need to write as much. Let us solve it again, but showing only what we need to.

**Solution.**

• We use integration by parts to solve the integral.

• Set $u(x) = x$ and $v'(x) = \sin x$. Then $u'(x) = 1$ and $v(x) = -\cos x$, and

$$\int x \sin x \,dx = -x \cos x + \int \cos x \,dx$$

$$= -x \cos x + \sin x + C.$$
It is pretty standard practice to reduce the notation even further in these problems. As noted above, many people write the integration by parts formula as

\[ \int u \, dv = uv - \int v \, du \]

where \( du, dv \) are shorthand for \( u'(x) \, dx, v'(x) \, dx \). Let us write up the previous example using this notation.

\[ \int x \sin x \, dx \]

**Solution.** Using integration by parts, we set \( u = x \) and \( dv = \sin x \, dx \). This makes \( du = 1 \, dx \) and \( v = -\cos x \). Consequently

\[ \int x \sin x \, dx = \int u \, dv = uv - \int v \, du = -x \cos x + \int \cos x \, dx = -x \cos x + \sin x + C \]

You can see that this is a very neat way to write up these problems and we will continue using this shorthand in the examples that follow below.

We can also use integration by parts to eliminate higher powers of \( x \). We just need to apply the method more than once.

\[ \int x^2 e^x \, dx \]

**Solution.**

- Let \( u = x^2 \) and \( dv = e^x \, dx \). This then gives \( du = 2x \, dx \) and \( v = e^x \), and

\[ \int x^2 e^x \, dx = x^2 e^x - \int 2x e^x \, dx \]

- So we have reduced the problem of computing the original integral to one of integrating \( 2xe^x \). We know how to do this — just integrate by parts again:

\[ \int 2xe^x \, dx = x^2 e^x - \int 2e^x \, dx \]

\[ = x^2 e^x - \left( 2e^x - \int 2e^x \, dx \right) \]

\[ = x^2 e^x - 2e^x + 2e^x + C \]

\[ = x^2 e^x + C \]
• We can, if needed, check our answer by differentiating:

\[
\frac{d}{dx} \left( x^2e^x - 2xe^x + 2e^x + C \right) = (x^2e^x + 2xe^x) - (2xe^x + 2e^x) + 2e^x + 0 = x^2e^x
\]

A similar iterated application of integration by parts will work for integrals

\[
\int P(x) \left( Ae^{ax} + B \sin(bx) + C \cos(cx) \right) \, dx
\]

where \( P(x) \) is a polynomial and \( A, B, C, a, b, c \) are constants.

Example 1.7.6

Now let us look at integrands containing logarithms. We don’t know the antiderivative of \( \log x \), but we can eliminate \( \log x \) from an integrand by using integration by parts with \( u = \log x \). Remember \( \log x = \log e \cdot x = \ln x \).

Example 1.7.7 (\( \int x \log x \, dx \))

Solution.

• We have two choices for \( u \) and \( dv \).

(1) Set \( u = x \) and \( dv = \log x \, dx \). This gives \( du = dx \) but \( v \) is hard to compute — we haven’t done it yet\(^{49}\). Before we go further along this path, we should look to see what happens with the other choice.

(2) Set \( u = \log x \) and \( dv = x \, dx \). This gives \( du = \frac{1}{x} \, dx \) and \( v = \frac{1}{2}x^2 \), and we have to integrate

\[
\int v \, du = \int \frac{1}{x} \cdot \frac{1}{2}x^2 \, dx
\]

which is easy.

• So we proceed with the second choice.

\[
\int x \log x \, dx = \frac{1}{2}x^2 \log x - \int \frac{1}{2}x \, dx
\]

\[
= \frac{1}{2}x^2 \log x - \frac{1}{4}x^2 + C
\]

• We can check our answer quickly:

\[
\frac{d}{dx} \left( \frac{x^2}{2} \ln x - \frac{x^2}{4} + C \right) = x \ln x + \frac{x^2}{2} \cdot \frac{1}{x} - \frac{x}{2} + 0 = x \ln x
\]

\(^{49}\) We will soon.
Example 1.7.7

It is not immediately obvious that one should use integration by parts to compute the integral

$$\int \log x \, dx$$

since the integrand is not a product. But we should persevere — indeed this is a situation where our shorter notation helps to clarify how to proceed.

**Solution.**

- In the previous example we saw that we could remove the factor $\log x$ by setting $u = \log x$ and using integration by parts. Let us try repeating this. When we make this choice, we are then forced to take $dv = dx$ — that is we choose $v'(x) = 1$. Once we have made this sneaky move everything follows quite directly.

- We then have $du = \frac{1}{x} \, dx$ and $v = x$, and the integration by parts formula gives us

$$\int \log x \, dx = x \log x - \int \frac{1}{x} \cdot x \, dx$$

$$= x \log x - \int 1 \, dx$$

$$= x \log x - x + C$$

- As always, it is a good idea to check our result by verifying that the derivative of the answer really is the integrand.

$$\frac{d}{dx}(x \ln x - x + C) = \ln x + x \frac{1}{x} - 1 + 0 = \ln x$$

Example 1.7.8

The same method works almost exactly to compute the antiderivatives of $\arcsin(x)$ and $\arctan(x)$:

Example 1.7.9

Compute the antiderivatives of the inverse sine and inverse tangent functions.

**Solution.**

- Again neither of these integrands are products, but that is no impediment. In both cases we set $dv = dx$ (i.e. $v'(x) = 1$) and choose $v(x) = x$. 

---

107
• For inverse tan we choose $u = \arctan(x)$, so $du = \frac{1}{1+x^2}dx$:

$$\int \arctan(x)dx = x \arctan(x) - \int x \cdot \frac{1}{1+x^2}dx \quad \text{now use substitution rule}$$

$$= x \arctan(x) - \int \frac{w'(x)}{2} \cdot \frac{1}{w}dx \quad \text{with } w(x) = 1 + x^2, w'(x) = 2x$$

$$= x \arctan(x) - \frac{1}{2} \int \frac{1}{w}dw$$

$$= x \arctan(x) - \frac{1}{2} \log |w| + C$$

$$= x \arctan(x) - \frac{1}{2} \log |1 + x^2| + C \quad \text{but } 1 + x^2 > 0, \text{ so}$$

$$= x \arctan(x) - \frac{1}{2} \log(1 + x^2) + C$$

• Similarly for inverse sine we choose $u = \arcsin(x)$ so $du = \frac{1}{\sqrt{1-x^2}}dx$:

$$\int \arcsin(x)dx = x \arcsin(x) - \int \frac{x}{\sqrt{1-x^2}}dx \quad \text{now use substitution rule}$$

$$= x \arcsin(x) - \int -\frac{w'(x)}{2} \cdot w^{-1/2}dx \quad \text{with } w(x) = 1 - x^2, w'(x) = -2x$$

$$= x \arcsin(x) + \frac{1}{2} \int w^{-1/2}dw$$

$$= x \arcsin(x) + \frac{1}{2} \cdot 2w^{1/2} + C$$

$$= x \arcsin(x) + \sqrt{1 - x^2} + C$$

• Both can be checked quite quickly by differentiating — but we leave that as an exercise for the reader.

There are many other examples we could do, but we’ll finish with a tricky one.

Example 1.7.10 ($\int e^x \sin x \, dx$)

Solution. Let us attempt this one a little naively and then we’ll come back and do it more carefully (and successfully).

• We can choose either $u = e^x, dv = \sin x \, dx$ or the other way around.

1. Let $u = e^x, dv = \sin x \, dx$. Then $du = e^x \, dx$ and $v = -\cos x$. This gives

$$\int e^x \sin x = -e^x \cos x + \int e^x \cos x \, dx$$

So we are left with an integrand that is very similar to the one we started with. What about the other choice?
2. Let \( u = \sin x \), \( dv = e^x dx \). Then \( du = \cos x dx \) and \( v = e^x \). This gives

\[
\int e^x \sin x = e^x \sin x - \int e^x \cos x dx
\]

So we are again left with an integrand that is very similar to the one we started with.

- How do we proceed? — It turns out to be easier if you do both \( \int e^x \sin x dx \) and \( \int e^x \cos x dx \) simultaneously. We do so in the next example.

Example 1.7.10

This time we’re going to do the two integrals

\[
I_1 = \int_a^b e^x \sin x dx \quad I_2 = \int_a^b e^x \cos x dx
\]

at more or less the same time.

- First

\[
I_1 = \int_a^b e^x \sin x dx = \int_a^b udv \quad \text{with } u = e^x, \ dv = \sin x dx
\]

so \( v = - \cos x \), \( du = e^x dx \)

\[
= \left[ - e^x \cos x \right]_a^b + \int_a^b e^x \cos x dx
\]

We have not found \( I_1 \) but we have related it to \( I_2 \).

\[
I_1 = \left[ - e^x \cos x \right]_a^b + I_2
\]

- Now start over with \( I_2 \).

\[
I_2 = \int_a^b e^x \cos x dx = \int_a^b udv \quad \text{with } u = e^x, \ dv = \cos x dx
\]

so \( v = \sin x \), \( du = e^x dx \)

\[
= \left[ e^x \sin x \right]_a^b - \int_a^b e^x \sin x dx
\]

Once again, we have not found \( I_2 \) but we have related it back to \( I_1 \).

\[
I_2 = \left[ e^x \sin x \right]_a^b - I_1
\]
1.8 Trigonometric Integrals

Integrals of polynomials of the trigonometric functions \( \sin x \), \( \cos x \), \( \tan x \) and so on, are generally evaluated by using a combination of simple substitutions and trigonometric identities. There are of course a very large number\(^{50}\) of trigonometric identities, but usually we use only a handful of them. The most important three are:

\(^{50}\) The more pedantic reader could construct an infinite list of them.
\[ \sin^2 x + \cos^2 x = 1 \]  
\text{Equation 1.8.1.}

\[ \sin(2x) = 2 \sin x \cos x \]  
\text{Equation 1.8.2.}

\[ \cos(2x) = \cos^2 x - \sin^2 x \]
\[ = 2 \cos^2 x - 1 \]
\[ = 1 - 2 \sin^2 x \]  
\text{Equation 1.8.3.}

Notice that the last two lines of Equation (1.8.3) follow from the first line by replacing either \( \sin^2 x \) or \( \cos^2 x \) using Equation (1.8.1). It is also useful to rewrite these last two lines:

\[ \sin^2 x = \frac{1 - \cos(2x)}{2} \]  
\text{Equation 1.8.4.}

\[ \cos^2 x = \frac{1 + \cos(2x)}{2} \]  
\text{Equation 1.8.5.}

These last two are particularly useful since they allow us to rewrite higher powers of sine and cosine in terms of lower powers. For example:

\[
\sin^4(x) = \left(\frac{1 - \cos(2x)}{2}\right)^2 \\
= \frac{1}{4} - \frac{1}{2} \cos(2x) + \frac{1}{4} \cos^2(2x) \\
\text{by Equation (1.8.4)} \\
= \frac{1}{4} - \frac{1}{2} \cos(2x) + \frac{1}{8} (1 + \cos(4x)) \\
\text{use Equation (1.8.5)} \\
= \frac{3}{8} - \frac{1}{2} \cos(2x) + \frac{1}{8} \cos(4x)
\]
So while it was hard to integrate \( \sin^4(x) \) directly, the final expression is quite straightforward (with a little substitution rule).

There are many such tricks for integrating powers of trigonometric functions. Here we concentrate on two families

\[
\int \sin^m x \cos^n x \, dx \quad \text{and} \quad \int \tan^m x \sec^n x \, dx
\]

for integer \( n, m \). The details of the technique depend on the parity of \( n \) and \( m \) — that is, whether \( n \) and \( m \) are even or odd numbers.

### 1.8.1 Integrating \( \int \sin^m x \cos^n x \, dx \)

#### One of \( n \) and \( m \) is odd

Consider the integral \( \int \sin^2 x \cos x \, dx \). We can integrate this by substituting \( u = \sin x \) and \( du = \cos x \, dx \). This gives

\[
\int \sin^2 x \cos x \, dx = \int u^2 du = \frac{1}{3} u^3 + C = \frac{1}{3} \sin^3 x + C
\]

This method can be used whenever \( n \) is an odd integer.

- Substitute \( u = \sin x \) and \( du = \cos x \, dx \).
- This leaves an even power of cosines — convert them using \( \cos^2 x = 1 - \sin^2 x = 1 - u^2 \).

Here is an example.

**Example 1.8.6** \( \int \sin^2 x \cos^3 x \, dx \)

Start by factoring off one power of \( \cos x \) to combine with \( dx \) to get \( \cos x \, dx = du \).

\[
\int \sin^2 x \cos^3 x \, dx = \int \sin^2 x \cos x^2 \cos x \, dx = \int u^2 \cos x^2 \cos x \, dx = \int u^2 \cos x \, dx = \int u^2 (1 - u^2) du = \frac{u^3}{3} - \frac{u^5}{5} + C = \frac{\sin^3 x}{3} - \frac{\sin^5 x}{5} + C
\]

Of course if \( m \) is an odd integer we can use the same strategy with the roles of \( \sin x \) and \( \cos x \) exchanged. That is, we substitute \( u = \cos x \), \( du = -\sin x \, dx \) and \( \sin^2 x = 1 - \cos^2 x = 1 - u^2 \).
Both $n$ and $m$ are even

If $m$ and $n$ are both even, the strategy is to use the trig identities (1.8.4) and (1.8.5) to get back to the $m$ or $n$ odd case. This is typically more laborious than the previous case we studied. Here are a couple of examples that arise quite commonly in applications.

Example 1.8.7 ($\int \cos^2 x \, dx$)

By (1.8.5)

$$\int \cos^2 x \, dx = \frac{1}{2} \int [1 + \cos(2x)] \, dx = \frac{1}{2} \left[ x + \frac{1}{2} \sin(2x) \right] + C$$

Example 1.8.8 ($\int \cos^4 x \, dx$)

First we’ll prepare the integrand $\cos^4 x$ for easy integration by applying (1.8.5) a couple times. We have already used (1.8.5) once to get

$$\cos^2 x = \frac{1}{2} [1 + \cos(2x)]$$

Squaring it gives

$$\cos^4 x = \frac{1}{4} [1 + \cos(2x)]^2 = \frac{1}{4} + \frac{1}{2} \cos(2x) + \frac{1}{4} \cos^2(2x)$$

Now by (1.8.5) a second time

$$\cos^4 x = \frac{1}{4} + \frac{1}{2} \cos(2x) + \frac{1}{4} \left( \frac{1 + \cos(4x)}{2} \right) = \frac{3}{8} + \frac{1}{2} \cos(2x) + \frac{1}{8} \cos(4x)$$

Now it’s easy to integrate

$$\int \cos^4 x \, dx = \frac{3}{8} \int dx + \frac{1}{2} \int \cos(2x) \, dx + \frac{1}{8} \int \cos(4x) \, dx$$

$$= \frac{3}{8} x + \frac{1}{4} \sin(2x) + \frac{1}{32} \sin(4x) + C$$

Example 1.8.9 ($\int \cos^2 x \sin^2 x \, dx$)

Here we apply both (1.8.4) and (1.8.5).

$$\int \cos^2 x \sin^2 x \, dx = \frac{1}{4} \int [1 + \cos(2x)] [1 - \cos(2x)] \, dx$$

$$= \frac{1}{4} \int [1 - \cos^2(2x)] \, dx$$

113
We can then apply (1.8.5) again
\[
\int \frac{1}{4} \left[ 1 - \frac{1}{2} (1 + \cos(4x)) \right] dx \\
= \frac{1}{8} \int [1 - \cos(4x)] dx \\
= \frac{1}{8} x - \frac{1}{32} \sin(4x) + C
\]
Oof! We could also have done this one using (1.8.2) to write the integrand as \(\sin^2(2x)\) and then used (1.8.4) to write it in terms of \(\cos(4x)\).

Example 1.8.9

Example 1.8.10 \((\int_0^\pi \cos^2 x dx \text{ and } \int_0^\pi \sin^2 x dx)\)

Of course we can compute the definite integral \(\int_0^\pi \cos^2 x dx\) by using the antiderivative for \(\cos^2 x\) that we found in Example 1.8.7. But here is a trickier way to evaluate that integral, and also the integral \(\int_0^\pi \sin^2 x dx\) at the same time, very quickly without needing the antiderivative of Example 1.8.7.

Solution.

- Observe that \(\int_0^\pi \cos^2 x dx\) and \(\int_0^\pi \sin^2 x dx\) are equal because they represent the same area — look at the graphs below — the darkly shaded regions in the two graphs have the same area and the lightly shaded regions in the two graphs have the same area.

- Consequently,
\[
\int_0^\pi \cos^2 x dx = \int_0^\pi \sin^2 x dx = \frac{1}{2} \left[ \int_0^\pi \sin^2 x dx + \int_0^\pi \cos^2 x dx \right] \\
= \frac{1}{2} \int_0^\pi [\sin^2 x + \cos^2 x] dx \\
= \frac{1}{2} \int_0^\pi dx \\
= \frac{\pi}{2}
\]
1.8.2 Integrating $\int \tan^m x \sec^n x \, dx$

The strategy for dealing with these integrals is similar to the strategy that we used to evaluate integrals of the form $\int \sin^m x \cos^n x \, dx$ and again depends on the parity of the exponents $n$ and $m$. It uses

$$\frac{d}{dx} \tan x = \sec^2 x \quad \frac{d}{dx} \sec x = \sec x \tan x \quad 1 + \tan^2 x = \sec^2 x$$

We split the methods for integrating $\int \tan^m x \sec^n x \, dx$ into 5 cases which we list below. These will become much more clear after an example (or two).

1. When $m$ is odd and any $n$ — rewrite the integrand in terms of $\sin x$ and $\cos x$:

$$\tan^m x \sec^n x \, dx = \left( \frac{\sin x}{\cos x} \right)^m \left( \frac{1}{\cos x} \right)^n \, dx$$

and then substitute $u = \cos x$, $du = -\sin x \, dx$, $\sin^2 x = 1 - \cos^2 x = 1 - u^2$. See Examples 1.8.11 and 1.8.12.

2. Alternatively, if $m$ is odd and $n \geq 1$ move one factor of $\sec x \tan x$ to the side so that you can see $\sec x \tan x \, dx$ in the integral, and substitute $u = \sec x$, $du = \sec x \tan x \, dx$ and $\tan^2 x = \sec^2 x - 1 = u^2 - 1$. See Example 1.8.13.

3. If $n$ is even with $n \geq 2$, move one factor of $\sec^2 x$ to the side so that you can see $\sec^2 x \, dx$ in the integral, and substitute $u = \tan x$, $du = \sec^2 x \, dx$ and $\sec^2 x = 1 + \tan^2 x = 1 + u^2$. See Example 1.8.14.

4. When $m$ is even and $n = 0$ — that is the integrand is just an even power of tangent — we can still use the $u = \tan x$ substitution, after using $\tan^2 x = \sec^2 x - 1$ (possibly more than once) to create a $\sec^2 x$. See Example 1.8.16.

5. This leaves the case $n$ odd and $m$ even. There are strategies like those above for treating this case. But they are more complicated and also involve more tricks (that basically have to be memorized). Examples using them are provided in the optional section entitled “Integrating $\sec x$, $\csc x$, $\sec^3 x$ and $\csc^3 x$”, below. A more straightforward strategy uses another technique called “partial fractions”. We shall return to this strategy after we have learned about partial fractions. See Example 1.10.5 and 1.10.6 in Section 1.10.

51 You will need to memorise the derivatives of tangent and secant. However there is no need to memorise $1 + \tan^2 x = \sec^2 x$. To derive it very quickly just divide $\sin^2 x + \cos^2 x = 1$ by $\cos^2 x$. 

115
In this case we rewrite the integrand in terms of sine and cosine and then substitute \( u = \cos x \), \( du = -\sin x \, dx \).

**Example 1.8.11 \( \int \tan x \, dx \)**

**Solution.**

- Write the integrand \( \tan x = \frac{1}{\cos x} \sin x \).
- Now substitute \( u = \cos x \), \( du = -\sin x \, dx \) just as we did in treating integrands of the form \( \sin^m x \, \cos^n x \) with \( m \) odd.

\[
\int \tan x \, dx = \int \frac{1}{\cos x} \sin x \, dx \quad \text{substitute } u = \cos x
\]

\[
= \int \frac{1}{u} \cdot (-1) \, du
\]

\[
= -\log |u| + C
\]

\[
= -\log |\cos x| + C
\]

- can also write in terms of secant

\[
= \log |\cos x|^{-1} + C = \log |\sec x| + C
\]

**Example 1.8.12 \( \int \tan^3 x \, dx \)**

**Solution.**

- Write the integrand \( \tan^3 x = \frac{\sin^2 x}{\cos^3 x} \sin x \).
- Again substitute \( u = \cos x \), \( du = -\sin x \, dx \). We rewrite the remaining even powers of \( \sin x \) using \( \sin^2 x = 1 - \cos^2 x = 1 - u^2 \).
- Hence

\[
\int \tan^3 x \, dx = \int \frac{\sin^2 x}{\cos^3 x} \sin x \, dx \quad \text{substitute } u = \cos x
\]

\[
= \int \frac{1 - u^2}{u^3} \, (-1) \, du
\]

\[
= \frac{u^{-2}}{2} + \log |u| + C
\]

\[
= \frac{1}{2 \cos^2 x} + \log |\cos x| + C
\]

- can rewrite in terms of secant

\[
= \frac{1}{2 \sec^2 x} - \log |\sec x| + C
\]
**\( m \) is odd and \( n \geq 1 \) — odd power of tangent and at least one secant**

Here we collect a factor of \( \tan x \sec x \) and then substitute \( u = \sec x \) and \( du = \sec x \tan x \, dx \). We can then rewrite any remaining even powers of \( \tan x \) in terms of \( \sec x \) using \( \tan^2 x = \sec^2 x - 1 = u^2 - 1 \).

**Example 1.8.13 \( \int \tan^3 x \sec^4 x \, dx \)**

**Solution.**

- Start by factoring off one copy of \( \sec x \tan x \) and combine it with \( dx \) to form \( \sec x \tan x \, dx \), which will be \( du \).
- Now substitute \( u = \sec x \), \( du = \sec x \tan x \, dx \) and \( \tan^2 x = \sec^2 x - 1 = u^2 - 1 \).
- This gives

\[
\int \tan^3 x \sec^4 x \, dx = \int \tan^2 x \frac{\sec^3 x}{u^2-1} \, du
\]

\[
= \int [u^2 - 1]u^3 \, du
\]

\[
= \frac{u^6}{6} - \frac{u^4}{4} + C
\]

\[
= \frac{1}{6} \sec^6 x - \frac{1}{4} \sec^4 x + C
\]

**Example 1.8.14 \( \int \sec^4 x \, dx \)**

**Solution.**

- Factor off one copy of \( \sec^2 x \) and combine it with \( dx \) to form \( \sec^2 x \, dx \), which will be \( du \).
- Then substitute \( u = \tan x \), \( du = \sec^2 x \, dx \) and rewrite any remaining even powers of \( \sec x \) as powers of \( \tan x = u \) using \( \sec^2 x = 1 + \tan^2 x = 1 + u^2 \).
This gives

\[ \int \sec^4 x \, dx = \int \frac{\sec^2 x \, sec^2 x \, dx}{1 + u^2} \, du \]

\[ = \int [1 + u^2] \, du \]

\[ = u + \frac{u^3}{3} + C \]

\[ = \tan x + \frac{1}{3} \tan^3 x + C \]

**Example 1.8.14**

**Example 1.8.15** (\( \int \tan^3 x \sec^4 x \, dx \) — redux)

**Solution.** Let us revisit this example using this slightly different approach.

- Factor off one copy of \( \sec^2 x \) and combine it with \( dx \) to form \( \sec^2 x \, dx \), which will be \( du \).
- Then substitute \( u = \tan x \), \( du = \sec^2 x \, dx \) and rewrite any remaining even powers of \( \sec x \) as powers of \( \tan x = u \) using \( \sec^2 x = 1 + \tan^2 x = 1 + u^2 \).

This gives

\[ \int \tan^3 x \sec^4 x \, dx = \int \frac{\tan^3 x \sec^2 x \, sec^2 x \, dx}{u^3} \frac{1}{1 + u^2} \, du \]

\[ = \int [u^3 + u^5] \, du \]

\[ = \frac{u^4}{4} + \frac{u^6}{6} + C \]

\[ = \frac{1}{4} \tan^4 x + \frac{1}{6} \tan^6 x + C \]

- This is not quite the same as the answer we got above in Example 1.8.13. However we can show they are (nearly) equivalent. To do so we substitute \( v = \sec x \) and \( \tan^2 x = \sec^2 x - 1 = v^2 - 1 \):

\[ \frac{1}{6} \tan^6 x + \frac{1}{4} \tan^4 x = \frac{1}{6} (v^2 - 1)^3 + \frac{1}{4} (v^2 - 1)^2 \]

\[ = \frac{1}{6} (v^6 - 3v^4 + 3v^2 - 1) + \frac{1}{4} (v^4 - 2v^2 + 1) \]

\[ = \frac{v^6}{6} - \frac{v^4}{2} + \frac{v^2}{6} + \frac{v^4}{4} - \frac{v^2}{2} + \frac{1}{4} \]

\[ = \frac{v^6}{6} - \frac{v^4}{4} + v^2 + \left( \frac{1}{4} - \frac{1}{6} \right) \]

\[ = \frac{1}{6} \sec^6 x - \frac{1}{4} \sec^4 x + \frac{1}{12} \]
So while $\frac{1}{6} \tan^6 x + \frac{1}{4} \tan^4 x \neq \frac{1}{6} \sec^6 x - \frac{1}{4} \sec^4 x$, they only differ by a constant. Hence both are valid antiderivatives of $\tan^3 x \sec^4 x$.

$m$ is even and $n = 0$ — even powers of tangent

We integrate this by setting $u = \tan x$. For this to work we need to pull one factor of $\sec^2 x$ to one side to form $du = \sec^2 x \, dx$. To find this factor of $\sec^2 x$ we (perhaps repeatedly) apply the identity $\tan^2 x = \sec^2 x - 1$.

Example 1.8.16 ($\int \tan^4 x \, dx$)

**Solution.**

- There is no $\sec^2 x$ term present, so we try to create it from $\tan^4 x$ by using $\tan^2 x = \sec^2 x - 1$.

\[
\tan^4 x = \tan^2 x \cdot \tan^2 x = \tan^2 x \left[ \sec^2 x - 1 \right] = \tan^2 x \sec^2 x - \frac{\tan^2 x}{\sec^2 x - 1} = \tan^2 x \sec^2 x - \sec^2 x + 1
\]

- Now we can substitute $u = \tan x$, $du = \sec^2 x \, dx$.

\[
\int \tan^4 x \, dx = \int \frac{\tan^2 x}{u^2} \sec^2 x \, dx - \int \frac{\sec^2 x}{u} \, dx + \int dx = \int u^2 \, du - \int du + \int dx = \frac{u^3}{3} - u + x + C = \frac{\tan^3 x}{3} - \tan x + x + C
\]
• First pull out a factor of \( \tan^2 x \) to create a \( \sec^2 x \) factor:

\[
\tan^8 x = \tan^6 x \cdot \tan^2 x \\
= \tan^6 x \cdot [ \sec^2 x - 1 ] \\
= \tan^6 x \sec^2 x - \tan^6 x
\]

The first term is now ready to be integrated, but we need to reapply the method to the second term:

\[
= \tan^6 x \sec^2 x - \tan^4 x \cdot [ \sec^2 x - 1 ] \\
= \tan^6 x \sec^2 x - \tan^4 x \sec^2 x + \tan^4 x \\
do it again \\
= \tan^6 x \sec^2 x - \tan^4 x \sec^2 x + \tan^2 x \cdot [ \sec^2 x - 1 ] \\
= \tan^6 x \sec^2 x - \tan^4 x \sec^2 x + \tan^2 x \sec^2 x - \tan^2 x \\
and again \\
= \tan^6 x \sec^2 x - \tan^4 x \sec^2 x + \tan^2 x \sec^2 x - [ \sec^2 x - 1 ]
\]

• Hence

\[
\int \tan^8 x \, dx = \int [ \tan^6 x \sec^2 x - \tan^4 x \sec^2 x + \tan^2 x \sec^2 x - \sec^2 x + 1 ] \, dx \\
= \int [ \tan^6 x - \tan^4 x + \tan^2 x - 1 ] \sec^2 x \, dx + \int \, dx \\
= \int [ u^6 - u^4 + u^2 - 1 ] \, du + x + C \\
= \frac{u^7}{7} - \frac{u^5}{5} + \frac{u^3}{3} - u + x + C \\
= \frac{1}{7} \tan^7 x - \frac{1}{5} \tan^5 x + \frac{1}{3} \tan^3 x - \tan x + x + C
\]

Indeed this example suggests that for integer \( k \geq 0 \):

\[
\int \tan^{2k} x \, dx = \frac{1}{2k-1} \tan^{2k-1}(x) - \frac{1}{2k-3} \tan^{2k-3} x + \cdots - (-1)^k \tan x + (-1)^k x + C
\]

This last example also shows how we might integrate an odd power of tangent:

Example 1.8.17

Solution. We follow the same steps
• Pull out a factor of \( \tan^2 x \) to create a factor of \( \sec^2 x \):

\[
\tan^7 x = \tan^5 x \cdot \tan^2 x
\]
\[
= \tan^5 x \cdot [\sec^2 x - 1]
\]
\[
= \tan^5 x \sec^2 x - \tan^5 x
\]
do it again
\[
= \tan^5 x \sec^2 x - \tan^3 x \cdot [\sec^2 x - 1]
\]
\[
= \tan^5 x \sec^2 x - \tan^3 x \sec^2 x + \tan^3 x
\]
and again
\[
= \tan^5 x \sec^2 x - \tan^3 x \sec^2 x + \tan x [\sec^2 x - 1]
\]
\[
= \tan^5 x \sec^2 x - \tan^3 x \sec^2 x + \tan x \sec^2 x - \tan x
\]

• Now we can substitute \( u = \tan x \) and \( du = \sec^2 x \, dx \) and also use the result from Example 1.8.11 to take care of the last term:

\[
\int \tan^7 x \, dx = \int [\tan^5 x \sec^2 x - \tan^3 x \sec^2 x + \tan x \sec^2 x] \, dx - \int \tan x \, dx
\]

Now factor out the common \( \sec^2 x \) term and integrate \( \tan x \) via Example 1.8.11

\[
= \int [\tan^5 x - \tan^3 x + \tan x] \sec x \, dx - \log |\sec x| + C
\]
\[
= \int [u^5 - u^3 + u] \, du - \log |\sec x| + C
\]
\[
= \frac{u^6}{6} - \frac{u^4}{4} + \frac{u^2}{2} - \log |\sec x| + C
\]
\[
= \frac{1}{6} \tan^6 x - \frac{1}{4} \tan^4 x + \frac{1}{2} \tan^2 x - \log |\sec x| + C
\]

This example suggests that for integer \( k \geq 0 \):

\[
\int \tan^{2k+1} x \, dx = \frac{1}{2k} \tan^{2k}(x) - \frac{1}{2k-2} \tan^{2k-2} x + \ldots - (-1)^k \frac{1}{2} \tan^2 x + (-1)^k \log |\sec x| + C
\]

Example 1.8.18

Of course we have not considered integrals involving powers of \( \cot x \) and \( \csc x \). But they can be treated in much the same way as \( \tan x \) and \( \sec x \) were.

1.8.3 Optional — Integrating \( \sec x \), \( \csc x \), \( \sec^3 x \) and \( \csc^3 x \)

As noted above, when \( n \) is odd and \( m \) is even, one can use similar strategies as to the previous cases. However the computations are often more involved and more tricks need to be deployed. For this reason we make this section optional — the computations are definitely non-trivial. Rather than trying to construct a coherent “method” for this case, we instead give some examples to give the idea of what to expect.

Example 1.8.19 (\( \int \sec x \, dx \) — by trickery)

Solution. There is a very sneaky trick to compute this integral.
• The standard trick for this integral is to multiply the integrand by \(1 = \frac{\sec x + \tan x}{\sec x + \tan x}\)

\[
\sec x = \sec x \frac{\sec x + \tan x}{\sec x + \tan x} = \frac{\sec^2 x + \sec x \tan x}{\sec x + \tan x}
\]

• Notice now that the numerator of this expression is exactly the derivative its denominator. Hence we can substitute \(u = \sec x + \tan x\) and then \(du = (\sec x \tan x + \sec^2 x) \, dx\).

Hence

\[
\int \sec x \, dx = \int \sec x \frac{\sec x + \tan x}{\sec x + \tan x} \, dx = \int \frac{\sec^2 x + \sec x \tan x}{\sec x + \tan x} \, dx
\]

\[
= \int \frac{1}{u} \, du
\]

\[
= \log |u| + C
\]

\[
= \log |\sec x + \tan x| + C
\]

Example 1.8.19

There is a second method for integrating \(\int \sec x \, dx\), that is more tedious, but more straightforward. In particular, it does not involve a memorized trick. The integral \(\int \sec x \, dx\) is converted into the integral \(\int \frac{du}{1-u^2}\) by using the substitution \(u = \sin x\), \(du = \cos x \, dx\). The integral \(\int \frac{du}{1-u^2}\) is then integrated by the method of partial fractions, which we shall learn about in Section 1.10 “Partial Fractions”. The details are in Example 1.10.5 in those notes. This second method gives the answer

\[
\int \sec x \, dx = \frac{1}{2} \log \frac{1 + \sin x}{1 - \sin x} + C
\]

which appears to be different than the answer in Example 1.8.19. But they are really the same (of course) since

\[
\frac{1 + \sin x}{1 - \sin x} = \frac{(1 + \sin x)^2}{1 - \sin^2 x} = \frac{(1 + \sin x)^2}{\cos^2 x}
\]

\[
\Rightarrow \frac{1}{2} \log \frac{1 + \sin x}{1 - \sin x} = \frac{1}{2} \log \frac{(1 + \sin x)^2}{\cos^2 x} = \log \left| \frac{\sin x + 1}{\cos x} \right| = \log |\tan x + \sec x|
\]

Oof!

Example 1.8.20 (\(\int \csc x \, dx\) — by the \(u = \tan \frac{x}{2}\) substitution)

Solution. The integral \(\int \csc x \, dx\) may also be evaluated by both the methods above. That is either

• by multiplying the integrand by a cleverly chosen \(1 = \frac{\cot x - \csc x}{\cot x - \csc x}\) and then substituting \(u = \cot x - \csc x\), \(du = (-\csc^2 x + \csc x \cot x) \, dx\), or
INTEGRATION

1.8 TRIGONOMETRIC INTEGRALS

by substituting $u = \cos x$, $du = -\sin x \, dx$ to give $\int \csc x \, dx = -\int \frac{du}{1-u^2}$ and then using the method of partial fractions.

These two methods give the answers

$$\int \csc x \, dx = \log | \cot x - \csc x | + C = -\frac{1}{2} \log \frac{1+\cos x}{1-\cos x} + C \quad (1.8.1)$$

In this example, we shall evaluate $\int \csc x \, dx$ by yet a third method, which can be used to integrate rational functions\(^{52}\) of $\sin x$ and $\cos x$.

- This method uses the substitution

$$x = 2 \arctan u \quad \text{i.e.} \quad u = \tan \frac{x}{2} \quad \text{and} \quad dx = \frac{2}{1+u^2} \, du$$

— a half-angle substitution.

- To express $\sin x$ and $\cos x$ in terms of $u$, we first use the double angle trig identities (Equations 1.8.2 and 1.8.3 with $x \mapsto x/2$) to express $\sin x$ and $\cos x$ in terms of $\sin \frac{x}{2}$ and $\cos \frac{x}{2}$:

$$\sin x = 2 \sin \frac{x}{2} \cos \frac{x}{2}$$

$$\cos x = \cos^2 \frac{x}{2} - \sin^2 \frac{x}{2}$$

- We then use the triangle

$$\sqrt{1+u^2}$$

$$\frac{x}{2}$$

$$1$$

$$u$$

$\square$

to express $\sin \frac{x}{2}$ and $\cos \frac{x}{2}$ in terms of $u$. The bottom and right hand sides of the triangle have been chosen so that $\tan \frac{x}{2} = u$. This tells us that

$$\sin \frac{x}{2} = \frac{u}{\sqrt{1+u^2}}$$

$$\cos \frac{x}{2} = \frac{1}{\sqrt{1+u^2}}$$

- This in turn implies that:

$$\sin x = 2 \sin \frac{x}{2} \cos \frac{x}{2} = 2 \cdot \frac{u}{\sqrt{1+u^2}} \cdot \frac{1}{\sqrt{1+u^2}} = \frac{2u}{1+u^2}$$

$$\cos x = \cos^2 \frac{x}{2} - \sin^2 \frac{x}{2} = \frac{1}{1+u^2} - \frac{u^2}{1+u^2} = \frac{1-u^2}{1+u^2}$$

Oof!

\(52\) A rational function of $\sin x$ and $\cos x$ is a ratio with both the numerator and denominator being finite sums of terms of the form $a \sin^n x \cos^m x$, where $a$ is a constant and $m$ and $n$ are positive integers.
Let’s use this substitution to evaluate \( \int \csc x \, dx \).

\[
\int \csc x \, dx = \int \frac{1}{\sin x} \, dx = \int \frac{1 + u^2}{2u} \, 2 \, du
= \int \frac{1}{u} \, du = \log |u| + C
= \log \left| \tan \frac{x}{2} \right| + C
\]

To see that this answer is really the same as that in (1.8.1), note that

\[
\cot x - \csc x = \frac{\cos x - 1}{\sin x} = \frac{-2 \sin^2(x/2)}{2 \sin(x/2) \cos(x/2)} = -\tan \frac{x}{2}
\]

Example 1.8.20

Solution. The standard trick used to evaluate \( \int \sec^3 x \, dx \) is integration by parts.

- Set \( u = \sec x \), \( dv = \sec^2 x \, dx \). Hence \( du = \sec x \tan x \, dx \), \( v = \tan x \) and

\[
\int \sec^3 x \, dx = \int \frac{\sec x \, \sec^2 x \, dx}{u \, dv}
= \frac{\sec x \, \tan x}{u} - \int \frac{\sec x \, \tan x \, dx}{v \, du}
\]

- Since \( \tan^2 x + 1 = \sec^2 x \), we have \( \tan^2 x = \sec^2 x - 1 \) and

\[
\int \sec^3 x \, dx = \sec x \, \tan x - \int [\sec^3 x - \sec x] \, dx
= \sec x \, \tan x + \log |\sec x + \tan x| + C - \int \sec^3 x \, dx
\]

where we used \( \int \sec x \, dx = \log |\sec x + \tan x| + C \), which we saw in Example 1.8.19.

- Now moving the \( \int \sec^3 x \, dx \) from the right hand side to the left hand side

\[
2 \int \sec^3 x \, dx = \sec x \, \tan x + \log |\sec x + \tan x| + C \quad \text{and so}
\]

\[
\int \sec^3 x \, dx = \frac{1}{2} \sec x \, \tan x + \frac{1}{2} \log |\sec x + \tan x| + C
\]

for a new arbitrary constant \( C \) (which is just one half the old one).

Example 1.8.21

The integral \( \int \sec^3 x \, dx \) can also be evaluated by two other methods.
• Substitute \( u = \sin x, \ du = \cos x \, dx \) to convert \( \int \sec^3 x \, dx \) into \( \int \frac{du}{[1-u^2]^2} \) and evaluate the latter using the method of partial fractions. This is done in Example 1.10.6 in Section 1.10.

• Use the \( u = \tan \frac{x}{2} \) substitution. We use this method to evaluate \( \int \csc^3 x \, dx \) in Example 1.8.22, below.

**Example 1.8.22** \( (\int \csc^3 x \, dx – \text{by the } u = \tan \frac{x}{2} \text{ substitution}) \)

**Solution.** Let us use the half-angle substitution that we introduced in Example 1.8.20.

• In this method we set

\[
u = \tan \frac{x}{2}, \quad \frac{dx}{\sin^3 x} = \frac{2}{1+u^2} \, du \quad \sin x = \frac{2u}{1+u^2} \quad \cos x = \frac{1-u^2}{1+u^2}\]

• The integral then becomes

\[
\int \csc^3 x \, dx = \int \frac{1}{\sin^3 x} \, dx
= \int \left( \frac{1+u^2}{2u} \right)^3 \frac{2}{1+u^2} \, du
= \frac{1}{4} \int \frac{1+2u^2+u^4}{u^3} \, du
= \frac{1}{4} \left( \frac{u^{-2}}{-2} + 2 \log |u| + \frac{u^2}{2} \right) + C
= \frac{1}{8} \left( -\cot^2 \frac{x}{2} + 4 \log \left| \tan \frac{x}{2} \right| + \tan^2 \frac{x}{2} \right) + C
\]

Oof!

• This is a perfectly acceptable answer. But if you don’t like the \( \frac{x}{2} \)'s, they may be eliminated by using

\[
\tan^2 \frac{x}{2} - \cot^2 \frac{x}{2} = \frac{\sin^2 \frac{x}{2} - \cos^2 \frac{x}{2}}{\cos^2 \frac{x}{2} - \sin^2 \frac{x}{2}}
= \frac{\sin^4 \frac{x}{2} - \cos^4 \frac{x}{2}}{\sin^2 \frac{x}{2} \cos^2 \frac{x}{2}}
= \frac{( \sin^2 \frac{x}{2} - \cos^2 \frac{x}{2} ) ( \sin^2 \frac{x}{2} + \cos^2 \frac{x}{2} )}{\sin^2 \frac{x}{2} \cos^2 \frac{x}{2}}
= \frac{\sin^2 \frac{x}{2} - \cos^2 \frac{x}{2}}{\sin^2 \frac{x}{2} \cos^2 \frac{x}{2}} \quad \text{since } \sin^2 \frac{x}{2} + \cos^2 \frac{x}{2} = 1
= \frac{1}{4} \sin^2 x
\]

by (1.8.2) and (1.8.3)
and

\[
\tan \frac{x}{2} = \frac{\sin \frac{x}{2}}{\cos \frac{x}{2}} = \frac{\sin^2 \frac{x}{2}}{\sin \frac{x}{2} \cos \frac{x}{2}} = \frac{\frac{1}{2} \left[ 1 - \cos x \right]}{\frac{1}{2} \sin x}
\]

by (1.8.2) and (1.8.3)

So we may also write

\[
\int \csc^3 x \, dx = \frac{1}{2} \cot x \csc x + \frac{1}{2} \log |\csc x - \cot x| + C
\]

That last optional section was a little scary — let’s get back to something a little easier.

### 1.9 Trigonometric Substitution

In this section we discuss substitutions that simplify integrals containing square roots of the form

\[
\sqrt{a^2 - x^2} \quad \sqrt{a^2 + x^2} \quad \sqrt{x^2 - a^2}.
\]

When the integrand contains one of these square roots, then we can use trigonometric substitutions to eliminate them. That is, we substitute

\[
x = a \sin u \quad \text{or} \quad x = a \tan u \quad \text{or} \quad x = a \sec u
\]

and then use trigonometric identities

\[
\sin^2 \theta + \cos^2 \theta = 1 \quad \text{and} \quad 1 + \tan^2 \theta = \sec^2 \theta
\]

to simplify the result. To be more precise, we can

- eliminate \(\sqrt{a^2 - x^2}\) from an integrand by substituting \(x = a \sin u\) to give

  \[
  \sqrt{a^2 - x^2} = \sqrt{a^2 - a^2 \sin^2 u} = \sqrt{a^2 \cos^2 u} = |a \cos u|
  \]

- eliminate \(\sqrt{a^2 + x^2}\) from an integrand by substituting \(x = a \tan u\) to give

  \[
  \sqrt{a^2 + x^2} = \sqrt{a^2 + a^2 \tan^2 u} = \sqrt{a^2 \sec^2 u} = |a \sec u|
  \]

- eliminate \(\sqrt{x^2 - a^2}\) from an integrand by substituting \(x = a \sec u\) to give

  \[
  \sqrt{x^2 - a^2} = \sqrt{a^2 \sec^2 u - a^2} = \sqrt{a^2 \tan^2 u} = |a \tan u|
  \]
When we have used substitutions before, we usually gave the new integration variable, \( u \), as a function of the old integration variable \( x \). Here we are doing the reverse — we are giving the old integration variable, \( x \), in terms of the new integration variable \( u \). We may do so, as long as we may invert to get \( u \) as a function of \( x \). For example, with \( x = a \sin u \), we may take \( u = \arcsin \frac{x}{a} \). This is a good time for you to review the definitions of \( \arcsin \theta \), \( \arctan \theta \) and \( \text{arcsec} \theta \). See Section 2.12, “Inverse Functions”, of the CLP Mathematics 100 notes.

As a warm-up, consider the area of a quarter of the unit circle.

**Example 1.9.1 (Quarter of the unit circle)**

Compute the area of the unit circle lying in the first quadrant.

**Solution.** We know that the answer is \( \pi/4 \), but we can also compute this as an integral — we saw this way back in Example 1.1.16:

\[
\text{area} = \int_0^1 \sqrt{1 - x^2} \, dx
\]

- To simplify the integrand we substitute \( x = \sin u \). With this choice \( \frac{dx}{du} = \cos u \) and so \( dx = \cos u \, du \).

- We also need to translate the limits of integration and it is perhaps easiest to do this by writing \( u \) as a function of \( x \) — namely \( u(x) = \arcsin x \). Hence \( u(0) = 0 \) and \( u(1) = \pi/2 \).

- Hence the integral becomes

\[
\int_0^1 \sqrt{1 - x^2} \, dx = \int_0^{\pi/2} \sqrt{1 - \sin^2 u} \cdot \cos u \, du
\]

\[
= \int_0^{\pi/2} \cos^2 u \cdot \cos u \, du
\]

Notice that here we have used that the positive square root \( \sqrt{\cos^2 u} = |\cos u| = \cos u \) because \( \cos(u) \geq 0 \) for \( 0 \leq u \leq \pi/2 \).

- To go further we use the techniques of Section 1.8.

\[
\int_0^1 \sqrt{1 - x^2} \, dx = \int_0^{\pi/2} \cos^2 u \, du
\]

\[
= \frac{1}{2} \int_0^{\pi/2} (1 + \cos(2u)) \, du
\]

\[
= \frac{1}{2} \left[ \frac{u}{2} \frac{1}{2} \sin(2u) \right]_0^{\pi/2}
\]

\[
= \frac{1}{2} \left( \frac{\pi}{2} - 0 + \frac{\sin \pi}{2} - \frac{\sin 0}{2} \right)
\]

\[
= \frac{\pi}{4}
\]
Example 1.9.1

\[ \int \frac{x^2}{\sqrt{1-x^2}} \, dx \]

\text{Solution.} We proceed much as we did in the previous example.

- To simplify the integrand we substitute \( x = \sin u \). With this choice \( \frac{dx}{du} = \cos u \) and so \( dx = \cos u \, du \). Also note that \( u = \arcsin x \).

- The integral becomes
  \[
  \int \frac{x^2}{\sqrt{1-x^2}} \, dx = \int \frac{\sin^2 u}{\sqrt{1-\sin^2 u}} \cdot \cos u \, du \\
  = \int \frac{\sin^2 u}{\cos u} \cdot \cos u \, du \\
  = \int \sin^2 u \, du
  \]

- To proceed further we need to get rid of the square-root. Since \( u = \arcsin x \) has domain \(-1 \leq x \leq 1\) and range \(-\pi/2 \leq u \leq \pi/2\), it follows that \( \cos u \geq 0 \) (since cosine is non-negative on these inputs). Hence
  \[
  \sqrt{\cos^2 u} = \cos u \quad \text{when } -\pi/2 \leq u \leq \pi/2
  \]

- So our integral now becomes
  \[
  \int \frac{x^2}{\sqrt{1-x^2}} \, dx = \int \frac{\sin^2 u}{\cos u} \cdot \cos u \, du \\
  = \int \sin^2 u \, du \\
  = \frac{1}{2} \int (1 - \cos 2u) \, du \\
  = \frac{u}{2} - \frac{1}{4} \sin 2u + C \\
  = \frac{1}{2} \arcsin x - \frac{1}{4} \sin(2 \arcsin x) + C
  \]

- We can simplify this further using a double-angle identity. Recall that \( u = \arcsin x \) and that \( x = \sin u \). Then
  \[
  \sin 2u = 2 \sin u \cos u
  \]

  We can replace \( \cos u \) using \( \cos^2 u = 1 - \sin^2 u \). Taking a square-root of this formula gives \( \cos u = \pm \sqrt{1-\sin^2 u} \). We need the positive branch here since \( \cos u \geq 0 \) when \(-\pi/2 \leq u \leq \pi/2\) (which is exactly the range of \( \arcsin x \)). Continuing along:
  \[
  \sin 2u = 2 \sin u \cdot \sqrt{1-\sin^2 u} \\
  = 2x \sqrt{1-x^2}
  \]

128
Thus our solution is

\[ \int \frac{x^2}{\sqrt{1-x^2}} \, dx = \frac{1}{2} \arcsin x - \frac{1}{4} \sin(2 \arcsin x) + C \]
\[ = \frac{1}{2} \arcsin x - \frac{1}{2} x \sqrt{1-x^2} + C \]

The above two examples illustrate the main steps of the approach. The next example is similar, but with more complicated limits of integration.

Example 1.9.3 \( \left( \int_{a}^{r} \sqrt{r^2 - x^2} \, dx \right) \)

Let’s find the area of the shaded region in the sketch below.

We’ll set up the integral using vertical strips. The strip in the figure has width \( dx \) and height \( \sqrt{r^2 - x^2} \). So the area is given by the integral

\[ \text{area} = \int_{a}^{r} \sqrt{r^2 - x^2} \, dx \]

which is very similar to the previous example.

Solution.

- To evaluate the integral we substitute

  \[ x = x(u) = r \sin u \quad \text{dx} = \frac{dx}{du} \, du = r \cos u \, du \]

  It is also helpful to write \( u \) as a function of \( x \) — namely \( u = \arcsin \frac{x}{r} \).

- The integral runs from \( x = a \) to \( x = r \). These correspond to

  \[ u(r) = \arcsin \frac{r}{r} = \arcsin 1 = \frac{\pi}{2} \]
  \[ u(a) = \arcsin \frac{a}{r} \quad \text{which does not simplify further} \]
The integral then becomes
\[
\int_a^r \sqrt{r^2 - x^2} \, dx = \int_{\arcsin(a/r)}^{\pi/2} \sqrt{r^2 - r^2 \sin^2 u} \cdot r \cos u \, du
\]
\[
= \int_{\arcsin(a/r)}^{\pi/2} r^2 \sqrt{1 - \sin^2 u} \cdot \cos u \, du
\]
\[
= r^2 \int_{\arcsin(a/r)}^{\pi/2} \cos^2 u \cdot \cos u \, du
\]

To proceed further (as we did in Examples 1.9.1 and 1.9.2) we need to think about whether \( \cos u \) is positive or negative.

• Since \( a \) (as shown in the diagram) satisfies \( 0 \leq a \leq r \), we know that \( u(a) \) lies between \( \arcsin(0) = 0 \) and \( \arcsin(1) = \pi/2 \). Hence the variable \( u \) lies between 0 and \( \pi/2 \), and on this range \( \cos u \geq 0 \). This allows us get rid of the square-root:
\[
\sqrt{\cos^2 u} = |\cos u| = \cos u
\]

• Putting this fact into our integral we get
\[
\int_a^r \sqrt{r^2 - x^2} \, dx = r^2 \int_{\arcsin(a/r)}^{\pi/2} \cos^2 u \cdot \cos u \, du
\]
\[
= r^2 \int_{\arcsin(a/r)}^{\pi/2} \, du
\]
Recall the identity \( \cos^2 u = \frac{1+\cos 2u}{2} \) from Section 1.8
\[
= \frac{r^2}{2} \int_{\arcsin(a/r)}^{\pi/2} (1 + \cos 2u) \, du
\]
\[
= \frac{r^2}{2} \left[ u + \frac{1}{2} \sin(2u) \right]_{\arcsin(a/r)}^{\pi/2}
\]
\[
= \frac{r^2}{2} \left( \frac{\pi}{2} + \frac{1}{2} \sin \pi - \arcsin(a/r) - \frac{1}{2} \sin(2 \arcsin(a/r)) \right)
\]
\[
= \frac{r^2}{2} \left( \frac{\pi}{2} - \arcsin(a/r) - \frac{1}{2} \sin(2 \arcsin(a/r)) \right)
\]
Oof! But there is a little further to go before we are done.

• We can again simplify the term \( \sin(2 \arcsin(a/r)) \) using a double angle identity. Set \( \theta = \arcsin(a/r) \). Then \( \theta \) is the angle in the triangle on the right below. By the double angle formula for \( \sin(2\theta) \) (Equation (1.8.2))
\[
\sin(2\theta) = 2 \sin \theta \cos \theta
\]
\[
= 2 \frac{a}{r} \sqrt{r^2 - a^2} \frac{r}{\sqrt{r^2 - a^2}}.
\]
So finally the area is

\[
\text{area} = \int_a^r \sqrt{r^2 - x^2} \, dx
\]

\[
= \frac{r^2}{2} \left( \frac{\pi}{2} - \arcsin(a/r) - \frac{1}{2} \sin(2 \arcsin(a/r)) \right)
\]

\[
= \frac{\pi r^2}{4} - \frac{r^2}{2} \arcsin(a/r) - \frac{a}{2} \sqrt{r^2 - a^2}
\]

This is a relatively complicated formula, but we can make some “reasonableness” checks, by looking at special values of \( a \).

- If \( a = 0 \) the shaded region, in the figure at the beginning of this example, is exactly one quarter of a disk of radius \( r \) and so has area \( \frac{1}{4} \pi r^2 \). Substituting \( a = 0 \) into our answer does indeed give \( \frac{1}{4} \pi r^2 \).

- At the other extreme, if \( a = r \), the shaded region disappears completely and so has area 0. Subbing \( a = r \) into our answer does indeed give 0, since \( \arcsin 1 = \frac{\pi}{2} \).

Example 1.9.3

Example 1.9.4 \( \left( \int_a^r \sqrt{r^2 - x^2} \, dx \right) \)

The integral \( \int_a^r \sqrt{r^2 - x^2} \, dx \) looks a lot like the integral we just did in the previous 3 examples. It can also be evaluated using the trigonometric substitution \( x = r \sin u \) — but that is unnecessarily complicated. Just because you have now learned how to use trigonometric substitution doesn’t mean that you should forget everything you learned before.

Solution. This integral is much more easily evaluated using the simple substitution \( u = r^2 - x^2 \).

- Set \( u = r^2 - x^2 \). Then \( du = -2x \, dx \), and so

\[
\int_a^r x \sqrt{r^2 - x^2} \, dx = \int_{r^2 - a^2}^0 \sqrt{u} \, \frac{du}{-2}
\]

\[
= -\frac{1}{2} \left[ \frac{u^{3/2}}{3/2} \right]_{r^2 - a^2}^0
\]

\[
= \frac{1}{3} \left[ r^2 - a^2 \right]^{3/2}
\]

Example 1.9.4

Enough sines and cosines — let us try a tangent substitution.

---

To paraphrase the Law of the Instrument, possibly Mark Twain and definitely some psychologists, when you have a shiny new hammer, everything looks like a nail.
Example 1.9.5 \( \int \frac{dx}{x^2 \sqrt{9 + x^2}} \)

Solution. As per our guidelines at the start of this section, the presence of the square root term \( \sqrt{9 + x^2} \) tells us to substitute \( x = 3 \tan u \).

- Substitute
  \[ x = 3 \tan u \quad \text{and} \quad dx = 3 \sec^2 u \, du \]

This allows us to remove the square root:

\[
\sqrt{9 + x^2} = \sqrt{9 + 9 \tan^2 u} = 3\sqrt{1 + \tan^2 u} = 3\sqrt{\sec^2 u} = 3|\sec u|
\]

- Hence our integral becomes

\[
\int \frac{dx}{x^2 \sqrt{9 + x^2}} = \int \frac{3 \sec^2 u}{9 \tan^2 u \cdot 3|\sec u|} \, du
\]

- To remove the absolute value we must consider the range of values of \( u \) in the integral. Since \( x = 3 \tan u \) we have \( u = \arctan(x/3) \). The range of \( \arctan \) is \(-\pi/2 \leq \arctan \leq \pi/2\) and so \( u = \arctan(x/3) \) will always lie between \(-\pi/2\) and \(+\pi/2\). Hence \( \cos u \) will always be positive, which in turn implies that \( |\sec u| = \sec u \).

- Using this fact our integral becomes:

\[
\int \frac{dx}{x^2 \sqrt{9 + x^2}} = \int \frac{3 \sec^2 u}{27 \tan^2 u \cdot \sec u} \, du
\]

\[= \frac{1}{9} \int \frac{\sec u}{\tan^2 u} \, du \quad \text{since} \, \sec u > 0
\]

- Rewrite this in terms of sine and cosine

\[
\int \frac{dx}{x^2 \sqrt{9 + x^2}} = \frac{1}{9} \int \frac{\sec u}{\tan^2 u} \, du = \frac{1}{9} \int \frac{\cos u}{\sin^2 u} \, du = \frac{1}{9} \int \frac{\cos u}{\sin^2 u} \, du
\]

Now we can use the substitution rule with \( y = \sin u \) and \( dy = \cos u \, du \)

\[= \frac{1}{9} \int \frac{dy}{y^2} \quad \text{(1.9.3)}
\]

\[= -\frac{1}{9y} + C \quad \text{(1.9.4)}
\]

\[= -\frac{1}{9 \sin u} + C \quad \text{(1.9.5)}
\]

\( \text{To be pedantic, we mean the range of the “standard” arctangent function or its “principle value”. One can define other arctangent functions with different ranges.} \)
\begin{itemize}
  \item The original integral was a function of \( x \), so we still have to rewrite \( \sin u \) in terms of \( x \). Remember that \( x = 3 \tan u \) or \( u = \arctan(x/3) \). So \( u \) is the angle shown in the triangle below and we can read off the triangle that

\[
\sin u = \frac{x}{\sqrt{9 + x^2}}
\]

\[
\Rightarrow \int \frac{dx}{x^2\sqrt{9 + x^2}} = -\frac{\sqrt{9 + x^2}}{9x} + C
\]

\end{itemize}

\textbf{Example 1.9.5} \( \int \frac{x^2}{\sqrt{x^2 - 1}} \, dx \)

\textit{Solution.} This one requires a secant substitution, but otherwise is very similar to those above.

\begin{itemize}
  \item Set \( x = \sec u \) and \( dx = \sec u \tan u \, du \). Then

\[
\int \frac{x^2}{\sqrt{x^2 - 1}} \, dx = \int \frac{\sec^2 u}{\sqrt{\sec^2 u - 1}} \, \sec u \tan u \, du
\]

\[
= \int \sec^3 u \cdot \sec u \tan u \, du
\]

\[
= \int \sec^3 u \cdot \frac{\tan u}{\sqrt{\tan^2 u}} \, du
\]

\[
= \int \sec^3 u \, du
\]

since \( \tan^2 u = \sec^2 u - 1 \)

\end{itemize}

\item As before we need to consider the range of \( u \) values in order to determine the sign of \( \tan u \). Notice that the integrand is only defined when either \( x < -1 \) or \( x > 1 \); thus we should treat the cases \( x < -1 \) and \( x > 1 \) separately. Let us assume that \( x > 1 \) and we will come back to the case \( x < -1 \) at the end of the example.

When \( x > 1 \), the standard \( u = \arcsin x \) takes values in \((0, \pi/2)\). This follows since when \( 0 < \sin u < \pi/2 \), we have \( 0 < \cos u < 1 \) and so \( \sec u > 1 \). Further, when \( 0 < u < \pi/2 \), we have \( \tan u > 0 \). Thus \( |\tan u| = \tan u \).

\begin{itemize}
  \item Back to our integral:

\[
\int \frac{x^2}{\sqrt{x^2 - 1}} \, dx = \int \sec^3 u \cdot \frac{\tan u}{|\tan u|} \, du
\]

\[
= \int \sec^3 u \, du
\]

since \( \tan u > 0 \)

\end{itemize}

This is exactly Example 1.8.21

\[
= \frac{1}{2} \sec u \tan u + \frac{1}{2} \log |\sec u + \tan u| + C
\]
Since we started with a function of \( x \) we need to finish with one. We know that
\[
\sec u = x
\]
and then we can use trig identities
\[
\tan^2 u = \sec^2 u - 1 = x^2 - 1 \quad \text{so} \quad \tan u = \pm \sqrt{x^2 - 1} \quad \text{but we know} \quad \tan u \geq 0,
\]
\[
\tan u = \sqrt{x^2 - 1}
\]
Thus
\[
\int \frac{x^2}{\sqrt{x^2 - 1}} \, dx = \frac{1}{2} x \sqrt{x^2 - 1} + \frac{1}{2} \log |x + \sqrt{x^2 - 1}| + C
\]

The above holds when \( x > 1 \). We can confirm that it is also true when \( x < -1 \) by showing the right-hand side is a valid antiderivative of the integrand. To do so we must differentiate our answer. Notice that we do not need to consider the sign of \( x + \sqrt{x^2 - 1} \) when we differentiate since we have already seen that
\[
\frac{d}{dx} \log |x| = \frac{1}{x}
\]
when either \( x < 0 \) or \( x > 0 \). So the computation that follows applies to both \( x > 1 \) and \( x < -1 \). The expressions become quite long so we differentiate each term separately:
\[
\frac{d}{dx} \left[ x \sqrt{x^2 - 1} \right] = \left[ \sqrt{x^2 - 1} + \frac{x^2}{\sqrt{x^2 - 1}} \right]
\]
\[
= \frac{1}{\sqrt{x^2 - 1}} \left[ (x^2 - 1) + x^2 \right]
\]
\[
\frac{d}{dx} \log \left| x + \sqrt{x^2 - 1} \right| = \frac{1}{x + \sqrt{x^2 - 1}} \cdot \left[ 1 + \frac{x}{\sqrt{x^2 - 1}} \right]
\]
\[
= \frac{1}{x + \sqrt{x^2 - 1}} \cdot \frac{x + \sqrt{x^2 - 1} \sqrt{x^2 - 1}}{\sqrt{x^2 - 1}}
\]
\[
= \frac{1}{\sqrt{x^2 - 1}}
\]
Putting things together then gives us
\[
\frac{d}{dx} \left[ \frac{1}{2} x \sqrt{x^2 - 1} + \frac{1}{2} \log |x + \sqrt{x^2 - 1}| + C \right] = \frac{1}{2 \sqrt{x^2 - 1}} \left[ (x^2 - 1) + x^2 + 1 \right] + 0
\]
\[
= \frac{x^2}{\sqrt{x^2 - 1}}
\]
This tells us that our answer for \( x > 1 \) is also valid when \( x < 1 \) and so
\[
\int \frac{x^2}{\sqrt{x^2 - 1}} \, dx = \frac{1}{2} x \sqrt{x^2 - 1} + \frac{1}{2} \log |x + \sqrt{x^2 - 1}| + C
\]
when \( x < -1 \) and when \( x > 1 \).
The method, as we have demonstrated it above, works when our integrand contains the square root of very specific families of quadratic polynomials. In fact, the same method works for more general quadratic polynomials — all we need to do is complete the square.\footnote{If you have not heard of “completing the square” don’t worry. It is not a difficult method and it will only take you a few moments to learn. It refers to rewriting a quadratic polynomial $P(x) = ax^2 + bx + c$ as $P(x) = a(x + d)^2 + e$ for new constants $d, e.$}

Example 1.9.7 \( \int_3^5 \frac{\sqrt{x^2 - 2x - 3}}{x - 1} \, dx \)

This time we have an integral with a square root in the integrand, but the argument of the square root, while a quadratic function of $x$, is not in one of the standard forms $\sqrt{a^2 - x^2}$, $\sqrt{a^2 + x^2}$, $\sqrt{x^2 - a^2}$. The reason that it is not in one of those forms is that the argument, $x^2 - 2x - 3$, contains a term, namely $-2x$ that is of degree one in $x$. So we try to manipulate it into one of the standard forms by completing the square.

Solution.

- We first rewrite the quadratic polynomial $x^2 - 2x - 3$ in the form $(x - a)^2 + b$ for some constants $a, b$. The easiest way to do this is to expand both expressions and compare coefficients of $x$:

$$x^2 - 2x - 3 = (x - a)^2 + b = (x^2 - 2ax + a^2) + b$$

So — if we choose $-2a = -2$ (so the coefficients of $x^1$ match) and $a^2 + b = -3$ (so the coefficients of $x^0$ match), then both expressions are equal. Hence we set $a = 1$ and $b = -4$. That is

$$x^2 - 2x - 3 = (x - 1)^2 - 4$$

Many of you may have seen this method when learning to sketch parabolas.

- Once this is done we can convert the square root of the integrand into a standard form by making the simple substitution $y = x - 1$. Here goes

$$\int_3^5 \frac{\sqrt{x^2 - 2x - 3}}{x - 1} \, dx = \int_3^5 \frac{\sqrt{(x - 1)^2 - 4}}{x - 1} \, dx$$

$$= \int_2^4 \frac{\sqrt{y^2 - 4}}{y} \, dy$$

with $y = x - 1, dy = dx$

$$= \int_0^{\pi/3} \frac{\sqrt{4 \sec^2 u - 4}}{2 \sec u} \, 2 \sec u \tan u \, du$$

with $y = 2 \sec u$

and $dy = 2 \sec u \tan u \, du$

Notice that we could also do this in fewer steps by setting $(x - 1) = 2 \sec u, dx = 2 \sec u \tan u \, du.$

\footnotetext[55]{If you have not heard of “completing the square” don’t worry. It is not a difficult method and it will only take you a few moments to learn. It refers to rewriting a quadratic polynomial $P(x) = ax^2 + bx + c$ as $P(x) = a(x + d)^2 + e$ for new constants $d, e.$}
To get the limits of integration we used that

- the value of \( u \) that corresponds to \( y = 2 \) obeys \( 2 = y = 2 \sec u = \frac{2}{\cos u} \) or \( \cos u = 1 \), so that \( u = 0 \) works and
- the value of \( u \) that corresponds to \( y = 4 \) obeys \( 4 = y = 2 \sec u = \frac{2}{\cos u} \) or \( \cos u = \frac{1}{2} \), so that \( u = \pi/3 \) works.

Now returning to the evaluation of the integral, we simplify and continue.

\[
\int_{3}^{5} \frac{\sqrt{x^2 - 2x - 3}}{x - 1} \, dx = \int_{0}^{\pi/3} 2\sqrt{\sec^2 u - 1} \tan u \, du
\]
\[
= 2 \int_{0}^{\pi/3} \tan^2 u \, du \quad \text{since } \sec^2 u = 1 + \tan^2 u
\]

In taking the square root of \( \sec^2 u - 1 = \tan^2 u \) we used that \( \tan u \geq 0 \) on the range \( 0 \leq u \leq \frac{\pi}{3} \).

\[
= 2 \int_{0}^{\pi/3} [\sec^2 u - 1] \, du \quad \text{since } \sec^2 u = 1 + \tan^2 u, \text{ again}
\]
\[
= 2 \left[ \tan u - \frac{u}{\pi/3} \right]_{0}^{\pi/3}
\]
\[
= 2 \left[ \sqrt{3} - \frac{\pi}{3} \right]
\]

---

1.10 Partial Fractions

Partial fractions is the name given to a technique of integration that may be used to integrate any rational function.\(^{56}\) We already know how to integrate some simple rational functions

\[
\int \frac{1}{x} \, dx = \log |x| + C
\]
\[
\int \frac{1}{1 + x^2} \, dx = \arctan(x) + C
\]

Combining these with the substitution rule, we can integrate similar but more complicated rational functions:

\[
\int \frac{1}{2x + 3} \, dx = \frac{1}{2} \log |2x + 3| + C
\]
\[
\int \frac{1}{3 + 4x^2} \, dx = \frac{1}{2\sqrt{3}} \arctan \left( \frac{2x}{\sqrt{3}} \right) + C
\]

By summing such terms together we can integrate yet more complicated forms

\[
\int \left[ x + \frac{1}{x + 1} + \frac{1}{x - 1} \right] \, dx = \frac{x^2}{2} + \log |x + 1| + \log |x - 1| + C
\]

\(^{56}\) Recall that a rational function is the ratio of two polynomials.
However we are not (typically) presented with a rational function nicely decomposed into neat little pieces. It is far more likely that the rational function will be written as the ratio of two polynomials. For example:

$$\int \frac{x^3 + x}{x^2 - 1} \, dx$$

In this specific example it is not hard to confirm that

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

and hence

$$\int \frac{x^3 + x}{x^2 - 1} \, dx = \int \left[ x + \frac{1}{x + 1} + \frac{1}{x - 1} \right] \, dx$$

$$= \frac{x^2}{2} + \log |x + 1| + \log |x - 1| + C$$

Of course going in this direction (from a sum of terms to a single rational function) is straightforward. To be useful we need to understand how to do this in reverse: decompose a given rational function into a sum of simpler pieces that we can integrate.

Suppose that \(N(x)\) and \(D(x)\) are polynomials. The basic strategy is to write \(\frac{N(x)}{D(x)}\) as a sum of very simple, easy to integrate rational functions, namely

(1) polynomials — we shall see below that these are needed when the degree\(^{57}\) of \(N(x)\) is equal to or strictly bigger than the degree of \(D(x)\), and

(2) rational functions of the particularly simple form \(\frac{A}{(ax + b)^n}\) and

(3) rational functions of the form \(\frac{Ax + B}{(ax^2 + bx + c)^m}\).

We already know how to integrate the first two forms, and we’ll see how to integrate the third form in the near future.

To begin to explore this method of decomposition, let us go back to the example we just saw

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

The technique that we will use is based on two observations:

(1) The denominators on the left-hand side of are the factors of the denominator \(x^2 - 1 = (x - 1)(x + 1)\) on the right-hand side.

---

\(^{57}\) The degree of a polynomial is the largest power of \(x\). For example, the degree of \(2x^3 + 4x^2 + 6x + 8\) is three.
(2) Use $P(x)$ to denote the polynomial on the left hand side, and then use $N(x)$ and $D(x)$ to denote the numerator and denominator of the right hand side. That is

$$P(x) = x \quad N(x) = x^3 + x \quad D(x) = x^2 - 1.$$  

Then the degree of $N(x)$ is the sum of the degrees of $P(x)$ and $D(x)$. This is because the highest degree term in $N(x)$ is $x^3$, which comes from multiplying $P(x)$ by $D(x)$, as we see in

$$x + \frac{1}{x + 1} + \frac{1}{x - 1} = \frac{p(x)}{(x + 1)(x - 1) + (x - 1) + (x + 1)} = \frac{x^3 + x}{x^2 - 1}.$$  

More generally, the presence of a polynomial on the left hand side is signalled on the right hand side by the fact that the degree of the numerator is at least as large as the degree of the denominator.

1.10.1 ▶ Partial fraction decomposition examples

Rather than writing up the technique — known as the partial fraction decomposition — in full generality, we will instead illustrate it through a sequence of examples.

\[\text{Example 1.10.1} \left( \int \frac{x - 3}{x^2 - 3x + 2} \, dx \right)\]

In this example, we integrate $\frac{N(x)}{D(x)} = \frac{x - 3}{x^2 - 3x + 2}$.

Solution.

- Step 1. We first check to see if a polynomial $P(x)$ is needed. To do so, we check to see if the degree of the numerator, $N(x)$, is strictly smaller than the degree of the denominator $D(x)$. In this example, the numerator, $x - 3$, has degree one and that is indeed strictly smaller than the degree of the denominator, $x^2 - 3x + 2$, which is two. In this case\(^{58}\), we do not need to extract a polynomial $P(x)$ and we move on to step 2.

- Step 2. The second step is to factor the denominator

$$x^2 - 3x + 2 = (x - 1)(x - 2)$$

In this example it is quite easy, but in future examples (and quite possibly in your homework, quizzes and exam) you will have to work harder to factor the denominator. In Appendix A.16 we have written up some simple tricks for factoring polynomials. We will illustrate them in Example 1.10.3 below.

- Step 3. The third step is to write $\frac{x - 3}{x^2 - 3x + 2}$ in the form

$$\frac{x - 3}{x^2 - 3x + 2} = \frac{A}{x - 1} + \frac{B}{x - 2}$$

\(^{58}\) We will soon get to an example (Example 1.10.2 in fact) in which the numerator degree is at least as large as the denominator degree — in that situation we have to extract a polynomial $P(x)$ before we can move on to step 2.
for some constants $A$ and $B$. More generally, if the denominator consists of $n$ different linear factors, then we decompose the ratio as

$$\text{rational function} = \frac{A_1}{\text{linear factor 1}} + \frac{A_2}{\text{linear factor 2}} + \cdots + \frac{A_n}{\text{linear factor n}}$$

To proceed we need to determine the values of the constants $A$, $B$ and there are several different methods to do so. Here are two methods

- **Step 3 – Algebra Method.** This approach has the benefit of being conceptually clearer and easier, but the downside is that it is more tedious.

To determine the values of the constants $A$, $B$, we put the right-hand side back over the common denominator $(x - 1)(x - 2)$.

$$\frac{x - 3}{x^2 - 3x + 2} = \frac{A}{x - 1} + \frac{B}{x - 2} = \frac{A(x - 2) + B(x - 1)}{(x - 1)(x - 2)}$$

The fraction on the far left is the same as the fraction on the far right if and only if their numerators are the same.

$$x - 3 = A(x - 2) + B(x - 1)$$

Write the right hand side as a polynomial in standard form (i.e. collect up all $x$ terms and all constant terms)

$$x - 3 = (A + B)x + (\text{constant terms})$$

For these two polynomials to be the same, the coefficient of $x$ on the left hand side and the coefficient of $x$ on the right hand side must be the same. Similarly the coefficients of $x^0$ (i.e. the constant terms) must match. This gives us a system of two equations.

$$A + B = 1$$
$$-2A - B = -3$$

in the two unknowns $A, B$. We can solve this system by

- using the first equation, namely $A + B = 1$, to determine $A$ in terms of $B$:

$$A = 1 - B$$

- Substituting this into the remaining equation eliminates the $A$ from second equation, leaving one equation in the one unknown $B$, which can then be solved for $B$:

$$-2A - B = -3$$
$$-2(1 - B) - B = -3$$
$$-2 + B = -3$$
so $B = -1$

59 That is, we take the decomposed form and sum it back together.
– Once we know $B$, we can substitute it back into $A = 1 - B$ to get $A$.

$$A = 1 - B = 1 - (-1) = 2$$

Hence

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

• **Step 3 – Sneaky Method.** This takes a little more work to understand, but it is the more efficient of the two methods.

We wish to find $A$ and $B$ for which

$$\frac{x - 3}{(x - 1)(x - 2)} = \frac{A}{x - 1} + \frac{B}{x - 2}$$

Note that the denominator on the left hand side has been written in factored form.

– To determine $A$, we multiply both sides of the equation by $A$’s denominator, which is $x - 1$,

$$\frac{x - 3}{x - 2} = A + \frac{(x - 1)B}{x - 2}$$

and then we completely eliminate $B$ from the equation by evaluating at $x = 1$. This value of $x$ is chosen to make $x - 1 = 0$.

$$\left.\frac{x - 3}{x - 2}\right|_{x=1} = A + \left.\frac{(x - 1)B}{x - 2}\right|_{x=1} \implies A = \frac{1 - 3}{1 - 2} = 2$$

– To determine $B$, we multiply both sides of the equation by $B$’s denominator, which is $x - 2$,

$$\frac{x - 3}{x - 1} = \frac{(x - 2)A}{x - 1} + B$$

and then we completely eliminate $A$ from the equation by evaluating at $x = 2$. This value of $x$ is chosen to make $x - 2 = 0$.

$$\left.\frac{x - 3}{x - 1}\right|_{x=2} = \left.\frac{(x - 2)A}{x - 1}\right|_{x=2} + B \implies B = \frac{2 - 3}{2 - 1} = -1$$

Hence we have (the thankfully consistent answer)

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

Notice that no matter which method we use to find the constants we can easily check our answer by summing the terms back together:

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

**Step 4.** The final step is to integrate.

$$\int \frac{x - 3}{x^2 - 3x + 2} \, dx = \int \frac{2}{x - 1} \, dx + \int \frac{-1}{x - 2} \, dx = 2 \log |x - 1| - \log |x - 2| + C$$
Perhaps the first thing that you notice is that this process takes quite a few steps. However no single step is all that complicated; it only takes practice. With that said, let’s do another, slightly more complicated, one.

Example 1.10.2

\[
\int \frac{3x^3 - 8x^2 + 4x - 1}{x^2 - 3x + 2} \, dx
\]

In this example, we integrate \( \frac{N(x)}{D(x)} = \frac{3x^3 - 8x^2 + 4x - 1}{x^2 - 3x + 2} \).

Solution.

• Step 1. We first check to see if the degree of the numerator \( N(x) \) is strictly smaller than the degree of the denominator \( D(x) \). In this example, the numerator, \( 3x^3 - 8x^2 + 4x - 1 \), has degree three and the denominator, \( x^2 - 3x + 2 \), has degree two. As \( 3 \geq 2 \), we have to implement the first step.

The goal of the first step is to write \( \frac{N(x)}{D(x)} \) in the form

\[
\frac{N(x)}{D(x)} = P(x) + \frac{R(x)}{D(x)}
\]

with \( P(x) \) being a polynomial and \( R(x) \) being a polynomial of degree strictly smaller than the degree of \( D(x) \). The right hand side is \( \frac{P(x)D(x) + R(x)}{D(x)} \), so we have to express the numerator in the form \( N(x) = P(x)D(x) + R(x) \), with \( P(x) \) and \( R(x) \) being polynomials and with the degree of \( R \) being strictly smaller than the degree of \( D \). \( P(x)D(x) \) is a sum of expressions of the form \( ax^nD(x) \). We want to pull as many expressions of this form as possible out of the numerator \( N(x) \), leaving only a low degree remainder \( R(x) \).

We do this using long division — the same long division you learned in school, but with the base 10 replaced by \( x \).

– We start by observing that to get from the highest degree term in the denominator (\( x^2 \)) to the highest degree term in the numerator (\( 3x^3 \)), we have to multiply it by \( 3x \). So we write,

\[
x^2 - 3x + 2 \Big| 3x^3 - 8x^2 + 4x - 1
\]

In the above expression, the denominator is on the left, the numerator is on the right and \( 3x \) is written above the highest order term of the numerator. Always put lower powers of \( x \) to the right of higher powers of \( x \) — this mirrors how you do long division with numbers; lower powers of ten sit to the right of lower powers of ten.

– Now we subtract \( 3x \) times the denominator, \( x^2 - 3x + 2 \), which is \( 3x^3 - 9x^2 + 6x \), from the numerator.

\[
\text{Example 1.10.1}
\]
\[
\frac{3x}{x^2 - 2x - 1} = 3x(x^2 - 3x + 2)
\]

- This has left a remainder of \(x^2 - 2x - 1\). To get from the highest degree term in the denominator \(x^2\) to the highest degree term in the remainder \(x^2\), we have to multiply by 1. So we write,

\[
\frac{3x}{x^2 - 2x - 1} = \frac{3x + 1}{3x^3 - 8x^2 + 4x - 1} - \frac{3x(x^2 - 3x + 2)}{x^2 - 3x + 2} \cdot \frac{x - 3}{x - 3}
\]

- Now we subtract 1 times the denominator, \(x^2 - 3x + 2\), which is \(x^2 - 3x + 2\), from the remainder.

\[
\frac{3x}{x^2 - 2x - 1} = \frac{3x + 1}{3x^3 - 8x^2 + 4x - 1} - \frac{3x(x^2 - 3x + 2)}{x^2 - 3x + 2} \cdot \frac{x - 3}{x - 3}
\]

- This leaves a remainder of \(x - 3\). Because the remainder has degree 1, which is smaller than the degree of the denominator (being degree 2), we stop.

- In this example, when we subtracted \(3x(x^2 - 3x + 2)\) and \(1(x^2 - 3x + 2)\) from \(3x - 8x^2 + 4x - 1\) we ended up with \(x - 3\). That is,

\[
3x^3 - 8x^2 + 4x - 1 - 3x(x^2 - 3x + 2) - 1(x^2 - 3x + 2) = x - 3
\]

or, collecting the two terms proportional to \((x^2 - 3x + 2)\)

\[
3x^3 - 8x^2 + 4x - 1 - (3x + 1)(x^2 - 3x + 2) = x - 3
\]

Moving the \((3x + 1)(x^2 - 3x + 2)\) to the right hand side and dividing the whole equation by \(x^2 - 3x + 2\) gives

\[
\frac{3x^3 - 8x^2 + 4x - 1}{x^2 - 3x + 2} = 3x + 1 + \frac{x - 3}{x^2 - 3x + 2}
\]

And we can easily check this expression just by summing the two terms on the right-hand side.

We have written the integrand in the form \(\frac{N(x)}{D(x)} = P(x) + \frac{R(x)}{D(x)}\), with the degree of \(R(x)\) strictly smaller than the degree of \(D(x)\), which is what we wanted. Observe that \(R(x)\) is the final remainder of the long division procedure and \(P(x)\) is at the top of the long division computation.
I NTEGRATION

1.10 P ARTIAL F RACTIONS

\[ \frac{3x + 1}{3x^3 - 8x^2 + 4x - 1} \quad P(x) \]
\[ \frac{3x^3 - 9x^2 + 6x}{x^2 - 2x - 1} \quad N(x) \]
\[ 3x \cdot D(x) \]
\[ \frac{x^2 - 3x + 2}{x - 3} \quad 1 \cdot D(x) \]
\[ R(x) = N(x) - (3x + 1)D(x) \]

This is the end of Step 1. Oof! You should definitely practice this step.

- **Step 2.** The second step is to factor the denominator

\[ x^2 - 3x + 2 = (x - 1)(x - 2) \]

We already did this in Example 1.10.1.

- **Step 3.** The third step is to write \( \frac{x - 3}{x^2 - 3x + 2} \) in the form

\[ \frac{x - 3}{x^2 - 3x + 2} = \frac{A}{x - 1} + \frac{B}{x - 2} \]

for some constants \( A \) and \( B \). We already did this in Example 1.10.1. We found \( A = 2 \) and \( B = -1 \).

- **Step 4.** The final step is to integrate.

\[
\int \frac{3x^3 - 8x^2 + 4x - 1}{x^2 - 3x + 2} \, dx = \int [3x + 1] \, dx + \int \frac{2}{x - 1} \, dx + \int \frac{-1}{x - 2} \, dx
\]

\[ = \frac{3}{2} x^2 + x + 2 \log |x - 1| - \log |x - 2| + C \]

You can see that the integration step is quite quick — almost all the work is in preparing the integrand.

Example 1.10.2

Here is a very solid example. It is quite long and the steps are involved. However please persist. No single step is too difficult.

Example 1.10.3 \( \left( \int \frac{x^4 + 5x^3 + 16x^2 + 26x + 22}{x^3 + 3x^2 + 7x + 5} \, dx \right) \)

In this example, we integrate \( \frac{N(x)}{D(x)} = \frac{x^4 + 5x^3 + 16x^2 + 26x + 22}{x^3 + 3x^2 + 7x + 5} \).

Solution.

- **Step 1.** Again, we start by comparing the degrees of the numerator and denominator. In this example, the numerator, \( x^4 + 5x^3 + 16x^2 + 26x + 22 \), has degree four and the
denominator, \( x^3 + 3x^2 + 7x + 5 \), has degree three. As \( 4 \geq 3 \), we must execute the first step, which is to write \( \frac{N(x)}{D(x)} \) in the form
\[
\frac{N(x)}{D(x)} = P(x) + \frac{R(x)}{D(x)}
\]
with \( P(x) \) being a polynomial and \( R(x) \) being a polynomial of degree strictly smaller than the degree of \( D(x) \). This step is accomplished by long division, just as we did in Example 1.10.3. We’ll go through the whole process in detail again.

Actually — before you read on ahead, please have a go at the long division. It is good practice.

- We start by observing that to get from the highest degree term in the denominator (\( x^3 \)) to the highest degree term in the numerator (\( x^4 \)), we have to multiply by \( x \). So we write,
\[
\frac{x}{x^3 + 3x^2 + 7x + 5} \quad \frac{x^4 + 5x^3 + 16x^2 + 26x + 22}{x^4 + 3x^3 + 7x^2 + 5x}
\]

- Now we subtract \( x \) times the denominator \( x^3 + 3x^2 + 7x + 5 \), which is \( x^4 + 3x^3 + 7x^2 + 5x \), from the numerator.
\[
\frac{x}{x^3 + 3x^2 + 7x + 5} \quad \frac{x^4 + 5x^3 + 16x^2 + 26x + 22}{x^4 + 3x^3 + 7x^2 + 5x} \quad \frac{2x^3 + 9x^2 + 21x + 22}{2x^3 + 9x^2 + 21x + 22}
\]

- The remainder was \( 2x^3 + 9x^2 + 21x + 22 \). To get from the highest degree term in the denominator (\( x^3 \)) to the highest degree term in the remainder (\( 2x^3 \)), we have to multiply by \( 2 \). So we write,
\[
\frac{x}{x^3 + 3x^2 + 7x + 5} \quad \frac{x^4 + 5x^3 + 16x^2 + 26x + 22}{x^4 + 3x^3 + 7x^2 + 5x} \quad \frac{2x^3 + 9x^2 + 21x + 22}{2x^3 + 9x^2 + 21x + 22}
\]

- Now we subtract \( 2 \) times the denominator \( x^3 + 3x^2 + 7x + 5 \), which is \( 2x^3 + 6x^2 + 14x + 10 \), from the remainder.
\[
\frac{x}{x^3 + 3x^2 + 7x + 5} \quad \frac{x^4 + 5x^3 + 16x^2 + 26x + 22}{x^4 + 3x^3 + 7x^2 + 5x} \quad \frac{2x^3 + 9x^2 + 21x + 22}{2x^3 + 9x^2 + 21x + 22}
\]

- This leaves a remainder of \( 3x^2 + 7x + 12 \). Because the remainder has degree 2, which is smaller than the degree of the denominator, which is 3, we stop.
In this example, when we subtracted \( x(x^3 + 3x^2 + 7x + 5) \) and \( 2(x^3 + 3x^2 + 7x + 5) \) from \( x^4 + 5x^3 + 16x^2 + 26x + 22 \) we ended up with \( 3x^2 + 7x + 12 \). That is,
\[
x^4 + 5x^3 + 16x^2 + 26x + 22 - x(x^3 + 3x^2 + 7x + 5) - 2(x^3 + 3x^2 + 7x + 5) = 3x^2 + 7x + 12
\]
or, collecting the two terms proportional to \( (x^3 + 3x^2 + 7x + 5) \)
\[
x^4 + 5x^3 + 16x^2 + 26x + 22 - (x + 2)(x^3 + 3x^2 + 7x + 5) = 3x^2 + 7x + 12
\]
Moving the \( (x + 2)(x^3 + 3x^2 + 7x + 5) \) to the right hand side and dividing the whole equation by \( x^3 + 3x^2 + 7x + 5 \) gives
\[
\frac{x^4 + 5x^3 + 16x^2 + 26x + 22}{x^3 + 3x^2 + 7x + 5} = x + 2 + \frac{3x^2 + 7x + 12}{x^3 + 3x^2 + 7x + 5}
\]
This is of the form \( \frac{N(x)}{D(x)} = P(x) + \frac{R(x)}{D(x)} \), with the degree of \( R(x) \) strictly smaller than the degree of \( D(x) \), which is what we wanted. Observe, once again, that \( R(x) \) is the final remainder of the long division procedure and \( P(x) \) is at the top of the long division computation.

\[
\begin{align*}
x^3 + 3x^2 + 7x + 5 & \quad \quad \quad P(x) \\
\frac{x^4 + 5x^3 + 16x^2 + 26x + 22}{x^3 + 3x^2 + 7x + 5} & \quad \quad \quad 2x^3 + 9x^2 + 21x + 22 \\
\frac{2x^3 + 6x^2 + 14x + 10}{3x^2 + 7x + 12} & \quad \quad \quad R(x)
\end{align*}
\]

- **Step 2.** The second step is to factor the denominator \( D(x) = x^3 + 3x^2 + 7x + 5 \). In the “real world” factorisation of polynomials is often very hard. Fortunately, this is not the “real world” and there is a trick available to help us find this factorisation. The reader should take some time to look at Appendix A.16 before proceeding.

  - The trick exploits the fact that most polynomials that appear in homework assignments and on tests have integer coefficients and some integer roots. Any integer root of a polynomial that has integer coefficients, like \( D(x) = x^3 + 3x^2 + 7x + 5 \), must divide the constant term of the polynomial exactly. Why this is true is explained in Appendix A.16.

  - So any integer root of \( x^3 + 3x^2 + 7x + 5 \) must divide 5 exactly. Thus the only integers which can be roots of \( D(x) \) are \( \pm 1 \) and \( \pm 5 \). Of course, not all of these give roots of the polynomial — in fact there is no guarantee that any of them will be. We have to test each one.

---

61 One does not typically think of mathematics assignments or exams as nice kind places...The polynomials that appear in the “real world” are not so forgiving. Nature, red in tooth and claw — to quote Tennyson inappropriately (especially when this author doesn’t know any other words from the poem).

62 Appendix A.16 contains several simple tricks for factoring polynomials. We recommend that you have a look at them. Honestly it is a great appendix and an awesome read.
To test if $+1$ is a root, we sub $x = 1$ into $D(x)$:

$$D(1) = 1^3 + 3(1)^2 + 7(1) + 5 = 16$$

As $D(1) \neq 0$, $1$ is not a root of $D(x)$.

To test if $-1$ is a root, we sub it into $D(x)$:

$$D(-1) = (-1)^3 + 3(-1)^2 + 7(-1) + 5 = -1 + 3 - 7 + 5 = 0$$

As $D(-1) = 0$, $-1$ is a root of $D(x)$. As $-1$ is a root of $D(x)$, $(x - (-1)) = (x + 1)$ must factor $D(x)$ exactly. We can factor the $(x + 1)$ out of $D(x) = x^3 + 3x^2 + 7x + 5$ by long division once again.

Dividing $D(x)$ by $(x + 1)$ gives:

$$x + 1 \left\{ \begin{array}{c} x^2 + 2x + 5 \\ x^2 + 7x + 5 \\ x^2 + 2x \\ 5x + 5 \\ 0 \end{array} \right\} \xrightarrow{-2x} x^2(x + 1) \xrightarrow{-5} 2(x + 1) \xrightarrow{5} 5(x + 1)$$

This time, when we subtracted $x^2(x + 1)$ and $2x(x + 1)$ and $5(x + 1)$ from $x^3 + 3x^2 + 7x + 5$ we ended up with 0 — as we knew would happen, because we knew that $x + 1$ divides $x^3 + 3x^2 + 7x + 5$ exactly. Hence

$$x^3 + 3x^2 + 7x + 5 = x^2(x + 1) + 2x(x + 1) - 5(x + 1) = 0$$

or

$$x^3 + 3x^2 + 7x + 5 = x^2(x + 1) + 2x(x + 1) + 5(x + 1)$$

or

$$x^3 + 3x^2 + 7x + 5 = (x^2 + 2x + 5)(x + 1)$$

It isn’t quite time to stop yet; we should attempt to factor the quadratic factor, $x^2 + 2x + 5$. We can use the quadratic formula$^{63}$ to find the roots of $x^2 + 2x + 5$:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-2 \pm \sqrt{4 - 20}}{2} = \frac{-2 \pm \sqrt{-16}}{2}$$

Since this expression contains the square root of a negative number the equation $x^2 + 2x + 5 = 0$ has no real solutions; without the use of complex numbers, $x^2 + 2x + 5$ cannot be factored.

---

$^{63}$ To be precise, the quadratic equation $ax^2 + bx + c = 0$ has solutions

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$  

The term $b^2 - 4ac$ is called the discriminant and it tells us about the number of solutions. If the discriminant is positive then there are two real solutions. When it is zero, there is a single solution. And if it is negative, there is no real solutions (you need complex numbers to say more than this).
We have reached the end of step 2. At this point we have
\[
\frac{x^4 + 5x^3 + 16x^2 + 26x + 22}{x^3 + 3x^2 + 7x + 5} = x + 2 + \frac{3x^2 + 7x + 12}{(x + 1)(x^2 + 2x + 5)}
\]

• **Step 3.** The third step is to write \(\frac{3x^2 + 7x + 12}{(x + 1)(x^2 + 2x + 5)}\) in the form
\[
\frac{3x^2 + 7x + 12}{(x + 1)(x^2 + 2x + 5)} = \frac{A}{x + 1} + \frac{Bx + C}{x^2 + 2x + 5}
\]

for some constants \(A, B, C\).

Note that the numerator, \(Bx + c\) of the second term on the right hand side is not just a constant. It is of degree one, which is exactly one smaller than the degree of the denominator, \(x^2 + 2x + 5\). More generally, if the denominator consists of \(n\) different linear factors and \(m\) different quadratic factors, then we decompose the ratio as

\[
\text{rational function} = \frac{A_1}{\text{linear factor 1}} + \frac{A_2}{\text{linear factor 2}} + \ldots + \frac{A_n}{\text{linear factor n}} + \frac{B_1x + C_1}{\text{quadratic factor 1}} + \frac{B_2x + C_2}{\text{quadratic factor 2}} + \ldots + \frac{B_mx + C_m}{\text{quadratic factor m}}
\]

To determine the values of the constants \(A, B, C\), we put the right hand side back over the common denominator \((x + 1)(x^2 + 2x + 5)\).

\[
\frac{3x^2 + 7x + 12}{(x + 1)(x^2 + 2x + 5)} = \frac{A}{x + 1} + \frac{Bx + C}{x^2 + 2x + 5} = \frac{A(x^2 + 2x + 5) + (Bx + C)(x + 1)}{(x + 1)(x^2 + 2x + 5)}
\]

The fraction on the far left is the same as the fraction on the far right if and only if their numerators are the same.

\[
3x^2 + 7x + 12 = A(x^2 + 2x + 5) + (Bx + C)(x + 1)
\]

Again, as in Example 1.10.1, there are a couple of different ways to determine the values of \(A, B, C\) from this equation.

• **Step 3 – Algebra Method.** The conceptually clearest procedure is to write the right hand side as a polynomial in standard form (i.e. collect up all \(x^2\) terms, all \(x\) terms and all constant terms)

\[
3x^2 + 7x + 12 = (A + B)x^2 + (2A + B + C)x + (5A + C)
\]

For these two polynomials to be the same, the coefficient of \(x^2\) on the left hand side and the coefficient of \(x^2\) on the right hand side must be the same. Similarly the coefficients of \(x^1\) must match and the coefficients of \(x^0\) must match.

This gives us a system of three equations

\[
A + B = 3 \quad 2A + B + C = 7 \quad 5A + C = 12
\]

in the three unknowns \(A, B, C\). We can solve this system by...
using the first equation, namely \( A + B = 3 \), to determine \( A \) in terms of \( B \): \( A = 3 - B \).

Substituting this into the remaining two equations eliminates the \( A \)'s from these two equations, leaving two equations in the two unknowns \( B \) and \( C \).

\[
\begin{align*}
A &= 3 - B \\
2A + B + C &= 7 \\
5A + C &= 12
\end{align*}
\]

\[
\begin{align*}
\Rightarrow & \quad 2(3 - B) + B + C = 7 \\
\Rightarrow & \quad 5(3 - B) + C = 12 \\
& \quad -B + C = 1 \\
& \quad -5B + C = -3
\end{align*}
\]

Now we can use the equation \( -B + C = 1 \), to determine \( B \) in terms of \( C \): \( B = C - 1 \).

Substituting this into the remaining equation eliminates the \( B \)'s leaving an equation in the one unknown \( C \), which is easy to solve.

\[
\begin{align*}
B &= C - 1 \\
-5B + C &= -3 \\
\Rightarrow & \quad -5(C - 1) + C = -3 \\
\Rightarrow & \quad -4C = -8
\end{align*}
\]

So \( C = 2 \), and then \( B = C - 1 = 1 \), and then \( A = 3 - B = 2 \). Hence

\[
\frac{3x^2 + 7x + 12}{(x + 1)(x^2 + 2x + 5)} = \frac{2}{x + 1} + \frac{x + 2}{x^2 + 2x + 5}
\]

- **Step 3 – Sneaky Method.** While the above method is transparent, it is rather tedious. It is arguably better to use the second, sneakier and more efficient, procedure. In order for

\[
3x^2 + 7x + 12 = A(x^2 + 2x + 5) + (Bx + C)(x + 1)
\]

the equation must hold for all values of \( x \).

In particular, it must be true for \( x = -1 \). When \( x = -1 \), the factor \((x + 1)\) multiplying \( Bx + C \) is exactly zero. So \( B \) and \( C \) disappear from the equation, leaving us with an easy equation to solve for \( A \):

\[
3x^2 + 7x + 12 \bigg|_{x=-1} = \left[ A(x^2 + 2x + 5) + (Bx + C)(x + 1) \right]_{x=-1} \implies 8 = 4A \implies A = 2
\]

Sub this value of \( A \) back in and simplify.

\[
3x^2 + 7x + 12 = 2(x^2 + 2x + 5) + (Bx + C)(x + 1)
\]

\[
x^2 + 3x + 2 = (Bx + C)(x + 1)
\]

Since \((x + 1)\) is a factor on the right hand side, it must also be a factor on the left hand side.

\[
(x + 2)(x + 1) = (Bx + C)(x + 1) \quad \Rightarrow \quad (x + 2) = (Bx + C) \quad \Rightarrow \quad B = 1, \ C = 2
\]
So again we find that
\[
\frac{3x^2 + 7x + 12}{(x+1)(x^2 + 2x + 5)} = \frac{2}{x+1} + \frac{x + 2}{x^2 + 2x + 5}
\]
Thus our integrand can be written as
\[
\frac{x^4 + 5x^3 + 16x^2 + 26x + 22}{x^3 + 3x^2 + 7x + 5} = x + 2 + \frac{2}{x+1} + \frac{x + 2}{x^2 + 2x + 5}.
\]

• Step 4. Now we can finally integrate! The first two pieces are easy.
\[
\int (x + 2) \, dx = \frac{1}{2} x^2 + 2x \quad \int \frac{2}{x+1} \, dx = 2 \log |x + 1|
\]
(We’re leaving the arbitrary constant to the end of the computation.)
The final piece is a little harder. The idea is to complete the square\textsuperscript{64} in the denominator
\[
\frac{x + 2}{x^2 + 2x + 5} = \frac{x + 2}{(x + 1)^2 + 4}
\]
and then make a change of variables to make the fraction look like \(\frac{ay + b}{y^2 + 1}\). In this case
\[
\frac{x + 2}{(x + 1)^2 + 4} = \frac{1}{4} \frac{x + 2}{\left(\frac{x+1}{2}\right)^2 + 1}
\]
so we make the change of variables \(y = \frac{x+1}{2}, \, dy = \frac{dx}{2}, \, x = 2y - 1, \, dx = 2 \, dy\)
\[
\int \frac{x + 2}{(x + 1)^2 + 4} \, dx = \frac{1}{4} \int \frac{x + 2}{\left(\frac{x+1}{2}\right)^2 + 1} \, dx
\]
\[
= \frac{1}{4} \int \frac{(2y - 1) + 2}{y^2 + 1} \, 2 \, dy
\]
\[
= \frac{1}{2} \int \frac{2y + 1}{y^2 + 1} \, dy
\]
\[
= \int \frac{y}{y^2 + 1} \, dy + \frac{1}{2} \int \frac{1}{y^2 + 1} \, dy
\]
\textsuperscript{64} This same idea arose in Section 1.9. Given a quadratic written as
\[
Q(x) = ax^2 + bx + c
\]
rewrite it as
\[
Q(x) = a(x + d)^2 + e.
\]
We can determine \(d\) and \(e\) by expanding and comparing coefficients of \(x\):
\[
a x^2 + bx + c = a(x^2 + 2dx + d^2) + e = ax^2 + 2adx + (e + ad^2)
\]
Hence \(d = b/2a\) and \(e = c - ad^2\).
Both integrals are easily evaluated, using the substitution \( u = y^2 + 1, \ du = 2y \, dy \) for the first.

\[
\int \frac{y}{y^2 + 1} \, dy = \int \frac{1}{u} \, \frac{du}{2} = \frac{1}{2} \log |u| = \frac{1}{2} \log(y^2 + 1) = \frac{1}{2} \log \left( \left( \frac{x + 1}{2} \right)^2 + 1 \right)
\]

\[
\frac{1}{2} \int \frac{1}{y^2 + 1} \, dy = \frac{1}{2} \arctan y = \frac{1}{2} \arctan \left( \frac{x + 1}{2} \right)
\]

That’s finally it. Putting all of the pieces together

\[
\int \frac{x^4 + 5x^3 + 16x^2 + 26x + 22}{x^3 + 3x^2 + 7x + 5} \, dx = \frac{1}{2} x^2 + 2x + 2 \log |x + 1|
\]

\[
+ \frac{1}{2} \log \left( \left( \frac{x + 1}{2} \right)^2 + 1 \right) + \frac{1}{2} \arctan \left( \frac{x + 1}{2} \right) + C
\]

The best thing after working through a few a nice long examples is to do another nice long example — it is excellent practice\(^\text{65}\). We recommend that the reader attempt the problem before reading through our solution.

In this example, we integrate \( \frac{N(x)}{D(x)} = \frac{4x^3 + 23x^2 + 45x + 27}{x^3 + 5x^2 + 8x + 4} \).

- **Step 1.** The degree of the numerator \( N(x) \) is equal to the degree of the denominator \( D(x) \), so the first step to write \( \frac{N(x)}{D(x)} \) in the form

\[
\frac{N(x)}{D(x)} = P(x) + \frac{R(x)}{D(x)}
\]

with \( P(x) \) being a polynomial (which should be of degree 0, i.e. just a constant) and \( R(x) \) being a polynomial of degree strictly smaller than the degree of \( D(x) \). By long division

\[
x^3 + 5x^2 + 8x + 4 \div \frac{4x^3 + 23x^2 + 45x + 27}{4x^3 + 20x^2 + 32x + 16}
\]

\[
= \frac{3x^2 + 13x + 11}{4}
\]

\(^\text{65}\) At the risk of quoting Nietzsche, “That which does not kill us makes us stronger.” Though this author always preferred the logically equivalent contrapositive — “That which does not make us stronger will kill us.” However no one is likely to be injured by practicing partial fractions or looking up quotes on wikipedia. Its also a good excuse to remind yourself of what a contrapositive is — though we will likely look at them again when we get to sequences and series.

150
so

\[
\frac{4x^3 + 23x^2 + 45x + 27}{x^3 + 5x^2 + 8x + 4} = \frac{4}{x^3 + 5x^2 + 8x + 4} + \frac{3x^2 + 13x + 11}{x^3 + 5x^2 + 8x + 4}
\]

- **Step 2.** The second step is to factorise \(D(x) = x^3 + 5x^2 + 8x + 4\).
  
  - To start, we’ll try and guess an integer root. Any integer root of \(D(x)\) must divide the constant term, 4, exactly. Only \(\pm 1, \pm 2, \pm 4\) can be integer roots of \(x^3 + 5x^2 + 8x + 4\).
  
  - We test to see if \(\pm 1\) are roots.
    
    \[
    D(1) = (1)^3 + 5(1)^2 + 8(1) + 4 = 0 \quad \Rightarrow \quad x = 1 \text{ is not a root}
    \]
    
    \[
    D(-1) = (-1)^3 + 5(-1)^2 + 8(-1) + 4 = 0 \quad \Rightarrow \quad x = -1 \text{ is a root}
    \]
    
    So \((x + 1)\) must divide \(x^3 + 5x^2 + 8x + 4\) exactly.
    
  - By long division

    \[
    \begin{array}{cccccccccc}
    & & & x^2 & + & 4x & + & 4 \\
    \hline
    x + 1 & | & x^3 & + & 5x^2 & + & 8x & + & 4 \\
    & & & x^3 & + & x^2 & | & & & \\
    & & & & -x^3 & -5x^2 & & & \\
    & & & & & 4x^2 & + & 8x & + & 4 \\
    & & & & & & 4x^2 & + & 4x & | & \\
    & & & & & & & -4x^2 & -8x & & \\
    & & & & & & & & 4x & + & 4 \\
    & & & & & & & & & 4x & + & 4 \\
    & & & & & & & & & & 0
    \end{array}
    \]
    
    so

    \[
    x^3 + 5x^2 + 8x + 4 = (x + 1)(x^2 + 4x + 4) = (x + 1)(x + 2)(x + 2)
    \]

  - Notice that we could have instead checked whether or not \(\pm 2\) are roots

    \[
    D(2) = (2)^3 + 5(2)^2 + 8(2) + 4 = 0 \quad \Rightarrow \quad x = 2 \text{ is not a root}
    \]
    
    \[
    D(-2) = (-2)^3 + 5(-2)^2 + 8(-2) + 4 = 0 \quad \Rightarrow \quad x = -2 \text{ is a root}
    \]
    
  We now know that both \(-1\) and \(-2\) are roots of \(x^3 + 5x^2 + 8x + 4\) and hence both \((x + 1)\) and \((x + 2)\) are factors of \(x^3 + 5x^2 + 8x + 4\). Because \(x^3 + 5x^2 + 8x + 4\) is of degree three and the coefficient of \(x^3\) is 1, we must have \(x^3 + 5x^2 + 8x + 4 = (x + 1)(x + 2)(x + a)\) for some constant \(a\). Multiplying out the right hand side shows that the constant term is 2\(a\). So \(2a = 4\) and \(a = 2\).

This is the end of step 2. We now know that

\[
\frac{4x^3 + 23x^2 + 45x + 27}{x^3 + 5x^2 + 8x + 4} = x + 4 + \frac{3x^2 + 13x + 11}{(x + 1)(x + 2)^2}
\]
• **Step 3.** The third step is to write \( \frac{3x^2 + 13x + 11}{(x+1)(x+2)^2} \) in the form

\[
\frac{3x^2 + 13x + 11}{(x+1)(x+2)^2} = \frac{A}{x+1} + \frac{B}{x+2} + \frac{C}{(x+2)^2}
\]

for some constants \( A, B \) and \( C \).

Note that there are two terms on the right hand arising from the factor \( (x+2)^2 \). One has denominator \( (x+2) \) and one has denominator \( (x+2)^2 \). More generally, for each factor \( (x+a)^n \) in the denominator of the rational function on the left hand side, we include

\[
\frac{A_1}{x+a} + \frac{A_2}{(x+a)^2} + \cdots + \frac{A_n}{(x+a)^n}
\]

in the partial fraction expansion on the right hand side.\(^{66}\)

To determine the values of the constants \( A, B, C \), we put the right hand side back over the common denominator \( (x+1)(x+2)^2 \):

\[
\frac{3x^2 + 13x + 11}{(x+1)(x+2)^2} = \frac{A}{x+1} + \frac{B}{x+2} + \frac{C}{(x+2)^2} = \frac{A(x+2)^2 + B(x+1)(x+2) + C(x+1)}{(x+1)(x+2)^2}
\]

The fraction on the far left is the same as the fraction on the far right if and only if their numerators are the same.

\[
3x^2 + 13x + 11 = A(x+2)^2 + B(x+1)(x+2) + C(x+1)
\]

As in the previous examples, there are a couple of different ways to determine the values of \( A, B \) and \( C \) from this equation.

• **Step 3 – Algebra Method.** The conceptually clearest procedure is to write the right hand side as a polynomial in standard form (i.e. collect up all \( x^2 \) terms, all \( x \) terms and all constant terms)

\[
3x^2 + 13x + 11 = (A + B)x^2 + (4A + 3B + C)x + (4A + 2B + C)
\]

For these two polynomials to be the same, the coefficient of \( x^2 \) on the left hand side and the coefficient of \( x^2 \) on the right hand side must be the same. Similarly the coefficients of \( x^1 \) and the coefficients of \( x^0 \) (i.e. the constant terms) must match. This gives us a system of three equations,

\[
A + B = 3 \quad 4A + 3B + C = 13 \quad 4A + 2B + C = 11
\]

in the three unknowns \( A, B, C \). We can solve this system by

---

\( ^{66} \) This is justified in the (optional) subsection “Justification of the Partial Fraction Decompositions” below.
- using the first equation, namely \( A + B = 3 \), to determine \( A \) in terms of \( B \):

\[
A = 3 - B.
\]

- Substituting this into the remaining equations eliminates the \( A \), leaving two equations in the two unknown \( B, C \).

\[
4(3 - B) + 3B + C = 13 \quad 4(3 - B) + 2B + C = 11
\]

or

\[
-B + C = 1 \quad -2B + C = -1
\]

- We can now solve the first of these equations, namely \(-B + C = 1\), for \( B \) in terms of \( C \), giving \( B = C - 1 \).

- Substituting this into the last equation, namely \(-2B + C = -1\), gives \(-2(C - 1) + C = -1\) which is easily solved to give

\[
C = 3, \text{ and then } B = C - 1 = 2 \text{ and then } A = 3 - B = 1.
\]

Hence

\[
\frac{4x^3 + 23x^2 + 45x + 27}{x^3 + 5x^2 + 8x + 4} = 4 + \frac{3x^2 + 13x + 11}{(x + 1)(x + 2)^2} = 4 + \frac{1}{x + 1} + \frac{2}{x + 2} + \frac{3}{(x + 2)^2}
\]

- **Step 3 – Sneaky Method.** The second, sneakier, method for finding \( A, B \) and \( C \) exploits the fact that \( 3x^2 + 13x + 11 = A(x + 2)^2 + B(x + 1)(x + 2) + C(x + 1) \) must be true for all values of \( x \). In particular, it must be true for \( x = -1 \). When \( x = -1 \), the factor \((x + 1)\) multiplying \( B \) and \( C \) is exactly zero. So \( B \) and \( C \) disappear from the equation, leaving us with an easy equation to solve for \( A \):

\[
3x^2 + 13x + 11 \bigg|_{x=-1} = \left[A(x + 2)^2 + B(x + 1)(x + 2) + C(x + 1)\right]_{x=-1}
\]

\[
\Rightarrow 1 = A
\]

Sub this value of \( A \) back in and simplify.

\[
3x^2 + 13x + 11 = (1)(x + 2)^2 + B(x + 1)(x + 2) + C(x + 1)
\]

\[
2x^2 + 9x + 7 = B(x + 1)(x + 2) + C(x + 1) = (xB + 2B + C)(x + 1)
\]

Since \((x + 1)\) is a factor on the right hand side, it must also be a factor on the left hand side.

\[
(2x + 7)(x + 1) = (xB + 2B + C)(x + 1) \quad \Rightarrow \quad (2x + 7) = (xB + 2B + C)
\]

For the coefficients of \( x \) to match, \( B \) must be 2. For the constant terms to match, \( 2B + C \) must be 7, so \( C \) must be 3. Hence we again have

\[
\frac{4x^3 + 23x^2 + 45x + 27}{x^3 + 5x^2 + 8x + 4} = 4 + \frac{3x^2 + 13x + 11}{(x + 1)(x + 2)^2} = 4 + \frac{1}{x + 1} + \frac{2}{x + 2} + \frac{3}{(x + 2)^2}
\]

153
• **Step 4.** The final step is to integrate

\[
\int \frac{4x^3 + 23x^2 + 45x + 27}{x^3 + 5x^2 + 8x + 4} \, dx = \int 4 \, dx + \int \frac{1}{x + 1} \, dx + \int \frac{2}{x + 2} \, dx + \int \frac{3}{(x + 2)^2} \, dx
\]

\[
= 4x + \log |x + 1| + 2\log |x + 2| - \frac{3}{x + 2} + C
\]

• **Example 1.10.4**

The method of partial fractions is not just confined to the problem of integrating rational functions. There are other integrals — such as \(\int \sec x \, dx\) and \(\int \sec^3 x \, dx\) — that can be transformed (via substitutions) into integrals of rational functions. We encountered both of these integrals in Sections 1.8 and 1.9 on trigonometric integrals and substitutions.

• **Example 1.10.5** \(\int \sec x \, dx\)

**Solution.** In this example, we integrate \(\sec x\). It is not yet clear what this integral has to do with partial fractions. To get to a partial fractions computation, we first make one of our old substitutions.

\[
\int \sec x \, dx = \int \frac{1}{\cos x} \, dx
\]

massage the expression a little

\[
= \int \frac{\cos x}{\cos^2 x} \, dx
\]

substitute \(u = \sin x, \, du = \cos x \, dx\)

\[
= -\int \frac{du}{u^2 - 1}
\]

and use \(\cos^2 x = 1 - \sin^2 x = 1 - u^2\)

So we now have to integrate \(\frac{1}{u^2 - 1}\), which is a rational function of \(u\), and so is perfect for partial fractions.

• **Step 1.** The degree of the numerator, 1, is zero, which is strictly smaller than the degree of the denominator, \(u^2 - 1\), which is two. So the first step is skipped.

• **Step 2.** The second step is to factor the denominator:

\[
u^2 - 1 = (u - 1)(u + 1)
\]

• **Step 3.** The third step is to write \(\frac{1}{u^2 - 1}\) in the form

\[
\frac{1}{u^2 - 1} = \frac{1}{(u - 1)(u + 1)} = \frac{A}{u - 1} + \frac{B}{u + 1}
\]

for some constants \(A\) and \(B\).

• **Step 3 – Sneaky Method.**
– Multiply through by the denominator to get
\[ 1 = A(u + 1) + B(u - 1) \]

This equation must be true for all \( u \).
– If we now set \( u = 1 \) then we eliminate \( B \) from the equation leaving us with
\[ 1 = 2A \]
so \( A = 1/2 \).
– Similarly, if we set \( u = -1 \) then we eliminate \( A \), leaving
\[ 1 = -2B \]
which implies \( B = -1/2 \).

We have now found that \( A = 1/2, B = -1/2 \), so
\[ \frac{1}{u^2 - 1} = \frac{1}{2} \left[ \frac{1}{u - 1} - \frac{1}{u + 1} \right]. \]

– It is always a good idea to check our work.
\[ \frac{1/2}{u - 1} + \frac{-1/2}{u + 1} = \frac{1/2(u + 1) - 1/2(u - 1)}{(u - 1)(u + 1)} = \frac{1}{(u - 1)(u + 1)} \]

• Step 4. The final step is to integrate.
\[
\int \sec x \, dx = -\int \frac{du}{u^2 - 1} \quad \text{after substitution}
\]
\[
= -\frac{1}{2} \int \frac{du}{u - 1} + \frac{1}{2} \int \frac{du}{u + 1} \quad \text{partial fractions}
\]
\[
= -\frac{1}{2} \log |u - 1| + \frac{1}{2} \log |u + 1| + C
\]
\[
= -\frac{1}{2} \log |\sin(x) - 1| + \frac{1}{2} \log |\sin(x) + 1| + C \quad \text{rearrange a little}
\]
\[
= \frac{1}{2} \log \left| \frac{1 + \sin x}{1 - \sin x} \right| + C
\]

Notice that since \(-1 \leq \sin x \leq 1\), we are free to drop the absolute values in the last line if we wish.

**Example 1.10.5**

Another example in the same spirit, though a touch harder. Again, we saw this problem in Section 1.8 and 1.9.

**Example 1.10.6** (\( \int \sec^3 x \, dx \))

**Solution.**
• We’ll start by converting it into the integral of a rational function using the substitution $u = \sin x$, $du = \cos x\,dx$.

\[
\int \sec^3 x\,dx = \int \frac{1}{\cos^3 x}\,dx = \int \frac{\cos x}{\cos^4 x}\,dx = \int \frac{\cos x}{(1 - \sin^2 x)^2}\,dx = \int \frac{du}{(1 - u^2)^2}
\]

massage this a little replace $\cos^2 x = 1 - \sin^2 x = 1 - u^2$

• We could now find the partial fractions expansion of the integrand $\frac{1}{(1 - u^2)^2}$ by executing the usual four steps. But it is easier to use

\[
\frac{1}{u^2 - 1} = \frac{1}{2} \left( \frac{1}{u - 1} - \frac{1}{u + 1} \right)
\]

which we worked out in Example 1.10.5 above.

• Squaring this gives

\[
\frac{1}{(1 - u^2)^2} = \frac{1}{4} \left[ \frac{1}{(u - 1)^2} - \frac{2}{(u - 1)(u + 1)} + \frac{1}{(u + 1)^2} \right]
\]

where we have again used $\frac{1}{u^2 - 1} = \frac{1}{2} \left( \frac{1}{u - 1} - \frac{1}{u + 1} \right)$ in the last step.

• It only remains to do the integrals and simplify.

\[
\int \sec^3 x\,dx = \frac{1}{4} \int \left[ \frac{1}{(u - 1)^2} - \frac{1}{u - 1} + \frac{1}{u + 1} + \frac{1}{(u + 1)^2} \right] du
\]

\[
= \frac{1}{4} \left[ -\frac{1}{u - 1} - \log |u - 1| + \log |u + 1| - \frac{1}{u + 1} \right] + C \quad \text{group carefully}
\]

\[
= -\frac{1}{4} \left[ \frac{1}{u - 1} + \frac{1}{u + 1} \right] + \frac{1}{4} \left[ \log |u + 1| - \log |u - 1| \right] + C \quad \text{sum carefully}
\]

\[
= -\frac{1}{4} \left[ \frac{2u}{u^2 - 1} + \frac{1}{u + 1} \right] + \frac{1}{4} \log \left| \frac{u + 1}{u - 1} \right| + C \quad \text{clean up}
\]

\[
= \frac{1}{2} \frac{u}{1 - u^2} + \frac{1}{4} \log \left| \frac{u + 1}{u - 1} \right| + C \quad \text{put } u = \sin x
\]

\[
= \frac{1}{2} \sin x + \frac{1}{4} \log \left| \frac{\sin x + 1}{\sin x - 1} \right| + C
\]

Example 1.10.6
1.10.2 The form of partial fractions decompositions

In the examples above we used the partial fractions method to decompose rational functions into easily integrated pieces. Each of those examples was quite involved and we had to spend quite a bit of time factoring and doing long division. The key step in each of the computations was Step 3 — in that step we decomposed the rational function \( \frac{N(x)}{D(x)} \) (or \( \frac{R(x)}{D(x)} \)), for which the degree of the numerator is strictly smaller than the degree of the denominator, into a sum of particularly simple rational functions, like \( \frac{A}{x-a} \). We did not, however, give a systematic description of those decompositions.

In this subsection we fill that gap by describing the general form of partial fraction decompositions. The justification of these forms is not part of the course, but the interested reader is invited to read the next (optional) subsection where such justification is given. In the following it is assumed that

- \( N(x) \) and \( D(x) \) are polynomials with the degree of \( N(x) \) strictly smaller than the degree of \( D(x) \).
- \( K \) is a constant.
- \( a_1, a_2, \cdots, a_j \) are all different numbers.
- \( m_1, m_2, \cdots, m_j \), and \( n_1, n_2, \cdots, n_k \) are all strictly positive integers.
- \( x^2 + b_1x + c_1, x^2 + b_2x + c_2, \cdots, x^2 + b_kx + c_k \) are all different.

### Simple linear factor case

If the denominator \( D(x) = K(x - a_1)(x - a_2) \cdots (x - a_j) \) is a product of \( j \) different linear factors, then

\[
\frac{N(x)}{D(x)} = \frac{A_1}{x - a_1} + \frac{A_2}{x - a_2} + \cdots + \frac{A_j}{x - a_j}
\]

Equation 1.10.7.

We can then integrate each term

\[
\int \frac{A}{x - a} \, dx = A \log |x - a| + C.
\]

### General linear factor case

If the denominator \( D(x) = K(x - a_1)^{m_1}(x - a_2)^{m_2} \cdots (x - a_j)^{m_j} \) then

Well — not the completely general form, in the sense that we are not allowing the use of complex numbers. As a result we have to use both linear and quadratic factors in the denominator. If we could use complex numbers we would be able to restrict ourselves to linear factors.
\[
\frac{N(x)}{D(x)} = \frac{A_{1,1}}{x-a_1} + \frac{A_{1,2}}{(x-a_1)^2} + \cdots + \frac{A_{1,m_1}}{(x-a_1)^{m_1}}
+ \frac{A_{2,1}}{x-a_2} + \frac{A_{2,2}}{(x-a_2)^2} + \cdots + \frac{A_{2,m_2}}{(x-a_2)^{m_2}} + \cdots
+ \frac{A_{j,1}}{x-a_j} + \frac{A_{j,2}}{(x-a_j)^2} + \cdots + \frac{A_{j,m_j}}{(x-a_j)^{m_j}}
\]

Notice that we could rewrite each line as
\[
\frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \cdots + \frac{A_m}{(x-a)^m} = A_1(x-a)^{m-1} + A_2(x-a)^{m-2} + \cdots + A_m
\]
\[
= B_1x^{m-1} + B_2x^{m-2} + \cdots + B_m
\]
which is a polynomial whose degree, \(m - 1\), is strictly smaller than that of the denominator \((x-a)^m\). But the form of Equation (1.10.8) is preferable because it is easier to integrate.

\[
\int \frac{A}{x-a} \, dx = A \log |x-a| + C
\]
\[
\int \frac{A}{(x-a)^k} \, dx = -\frac{1}{k-1} \cdot \frac{A}{(x-a)^{k-1}} \quad \text{provided } k > 1.
\]

**Simple linear and quadratic factor case**

If \(D(x) = K(x-a_1) \cdots (x-a_j)(x^2 + b_1x + c_1) \cdots (x^2 + b_kx + c_k)\) then

\[
\frac{N(x)}{D(x)} = \frac{A_1}{x-a_1} + \cdots + \frac{A_j}{x-a_j} + \frac{B_1x+C_1}{x^2+b_1x+c_1} + \cdots + \frac{B_kx+C_k}{x^2+b_kx+c_k}
\]

Note that the numerator of each term on the right hand side has degree one smaller than the degree of the denominator.

The quadratic terms \(\frac{Bx+C}{x^2+bx+c}\) are integrated in a two-step process that is best illustrated with a simple example (see also Example 1.10.3 above).

Example 1.10.10 \(\left( \int \frac{2x+7}{x^2+4x+13} \, dx \right)\)

Solution.

- Start by completing the square in the denominator:
  \[x^2 + 4x + 13 = (x + 2)^2 + 9\]
  and thus
  \[
  \frac{2x+7}{x^2+4x+13} = \frac{2x+7}{(x+2)^2+3^2}
  \]
1.10 Partial Fractions

- Now set $y = (x + 2)/3$, $dy = \frac{1}{3} dx$, or equivalently $x = 3y - 2$, $dx = 3dy$:

$$\int \frac{2x + 7}{x^2 + 4x + 13} dx = \int \frac{2x + 7}{(x + 2)^2 + 3^2} dx$$
$$= \int \frac{6y - 4 + 7}{3^2y^2 + 3^2} \cdot 3dy$$
$$= \int \frac{6y + 3}{3(y^2 + 1)} dy$$
$$= \int \frac{2y + 1}{y^2 + 1} dy$$

Notice that we chose 3 in $y = (x + 2)/3$ precisely to transform the denominator into the form $y^2 + 1$.

- Now almost always the numerator will be a linear polynomial of $y$ and we decompose as follows

$$\int \frac{2x + 7}{x^2 + 4x + 13} dx = \int \frac{2y + 1}{y^2 + 1} dy$$
$$= \int \frac{2y}{y^2 + 1} dy + \int \frac{1}{y^2 + 1} dy$$
$$= \log |y^2 + 1| + \arctan y + C$$
$$= \log \left| \left( \frac{x + 2}{3} \right)^2 + 1 \right| + \arctan \left( \frac{x + 2}{3} \right) + C$$

---

Optional — General linear and quadratic factor case

If $D(x) = K(x - a_1)^{m_1} \cdots (x - a_j)^{m_j} (x^2 + b_1 x + c_1)^{n_1} \cdots (x^2 + b_k x + c_k)^{n_k}$ then

**Equation 1.10.11.**
We have already seen how to integrate the simple and general linear terms, and the simple quadratic terms. Integrating general quadratic terms is not so straightforward.

**Example 1.10.12** \( \int \frac{dx}{(x^2 + 1)^n} \)

This example is not so easy, so it should definitely be considered optional.

**Solution.** In what follows write 

\[ I_n = \int \frac{dx}{(x^2 + 1)^n}. \]

- When \( n = 1 \) we know that 
  \[ \int \frac{dx}{x^2 + 1} = \arctan x + C \]

- Now assume that \( n > 1 \), then
  \[ \int \frac{1}{(x^2 + 1)^n} \, dx = \int \frac{(x^2 + 1 - x^2)}{(x^2 + 1)^n} \, dx \]
  \[ = \int \frac{1}{(x^2 + 1)^{n-1}} \, dx - \int \frac{x^2}{(x^2 + 1)^n} \, dx \]
  \[ = I_{n-1} - \int \frac{x^2}{(x^2 + 1)^n} \, dx \]

So we can write \( I_n \) in terms of \( I_{n-1} \) and this second integral.

- We can use integration by parts to compute the second integral:
  \[ \int \frac{x^2}{(x^2 + 1)^n} \, dx = \int \frac{x}{2} \cdot \frac{2x}{(x^2 + 1)^n} \, dx \]

We set \( u = x/2 \) and \( dv = \frac{2x}{(x^2 + 1)^n} \, dx \), which gives \( du = \frac{1}{2} \, dx \) and \( v = -\frac{1}{n-1} \cdot \frac{1}{(x^2 + 1)^{n-1}} \).

You can check \( v \) by differentiating. Integration by parts gives

\[ \int \frac{x}{2} \cdot \frac{2x}{(x^2 + 1)^n} \, dx = -\frac{x}{2(n-1)(x^2 + 1)^{n-1}} + \int \frac{dx}{2(n-1)(x^2 + 1)^{n-1}} \]

\[ = -\frac{x}{2(n-1)(x^2 + 1)^{n-1}} + \frac{1}{2(n-1)} \cdot I_{n-1} \]

- Now put everything together:
  \[ I_n = \frac{1}{(x^2 + 1)^n} \, dx \]
  \[ = I_{n-1} + \frac{x}{2(n-1)(x^2 + 1)^{n-1}} - \frac{1}{2(n-1)} \cdot I_{n-1} \]
  \[ = \frac{2n - 3}{2(n-1)} I_{n-1} + \frac{x}{2(n-1)(x^2 + 1)^{n-1}} \]
We can then use this recurrence to write down $I_n$ for the first few $n$:

\[
I_2 = \frac{1}{2} I_1 + \frac{x}{2(x^2 + 1)} + C
\]

\[
= \frac{1}{2} \arctan x + \frac{x}{2(x^2 + 1)}
\]

\[
I_3 = \frac{3}{4} I_2 + \frac{x}{4(x^2 + 1)^2}
\]

\[
= \frac{3}{8} \arctan x + \frac{3x}{8(x^2 + 1)} + \frac{x}{4(x^2 + 1)^2} + C
\]

\[
I_4 = \frac{5}{6} I_3 + \frac{x}{6(x^2 + 1)^3}
\]

\[
= \frac{5}{16} \arctan x + \frac{5x}{16(x^2 + 1)} + \frac{5x}{24(x^2 + 1)^2} + \frac{x}{6(x^2 + 1)^3} + C
\]

and so forth. You can see why partial fraction questions involving denominators with repeated quadratic factors do not often appear on exams.

---

1.10.3 Optional — Justification of the partial fraction decompositions

We will now see the justification for the form of the partial fraction decompositions. We will only consider the case in which the denominator has only linear factors. The arguments when there are quadratic factors too are similar.\(^{68}\)

>>> Simple linear factor case

In the most common partial fraction decomposition, we split up

\[
\frac{N(x)}{(x - a_1) \times \cdots \times (x - a_d)}
\]

into a sum of the form

\[
\frac{A_1}{x - a_1} + \cdots + \frac{A_d}{x - a_d}
\]

We now show that this decomposition can always be achieved, under the assumptions that the $a_i$'s are all different and $N(x)$ is a polynomial of degree at most $d - 1$. To do so, we shall repeatedly apply the following Lemma.

---

\(^{68}\) If we use complex numbers then every polynomial can be written as a product of linear factors. This is the fundamental theorem of algebra.
INTEGRATION

1.10 Partial Fractions

Let \( N(x) \) and \( D(x) \) be polynomials of degree \( n \) and \( d \) respectively, with \( n \leq d \). Suppose that \( a \) is NOT a zero of \( D(x) \). Then there is a polynomial \( P(x) \) of degree \( p < d \) and a number \( A \) such that

\[
\frac{N(x)}{D(x)} = \frac{P(x)}{D(x)} + \frac{A}{x-a}
\]

Proof:

• To save writing, let \( z = x - a \). We then write \( \tilde{N}(z) = N(z+a) \) and \( \tilde{D}(z) = D(z+a) \), which are again polynomials of degree \( n \) and \( d \) respectively. We also know that \( \tilde{D}(0) = D(a) \neq 0 \).

• In order to complete the proof we need to find a polynomial \( \tilde{P}(z) \) of degree \( p < d \) and a number \( A \) such that

\[
\frac{\tilde{N}(z)}{\tilde{D}(z)} = \frac{\tilde{P}(z)}{\tilde{D}(z)} + \frac{A}{z} = \frac{\tilde{P}(z)z + A\tilde{D}(z)}{\tilde{D}(z)}
\]

or equivalently, such that

\[
\tilde{P}(z)z + A\tilde{D}(z) = \tilde{N}(z).
\]

• Now look at the polynomial on the left hand side. Every term in \( \tilde{P}(z)z \), has at least one power of \( z \). So the constant term on the left hand side is exactly the constant term in \( A\tilde{D}(z) \), which is equal to \( A\tilde{D}(0) \). The constant term on the right hand side is equal to \( \tilde{N}(0) \). So the constant terms on the left and right hand sides are the same if we choose \( A = \frac{\tilde{N}(0)}{\tilde{D}(0)} \). Recall that \( \tilde{D}(0) \) cannot be zero, so \( A \) is well defined.

• Now move \( A\tilde{D}(z) \) to the right hand side.

\[
\tilde{P}(z)z = \tilde{N}(z) - A\tilde{D}(z)
\]

The constant terms in \( \tilde{N}(z) \) and \( A\tilde{D}(z) \) are the same, so the right hand side contains no constant term and the right hand side is of the form \( \tilde{N}_1(z)z \) for some polynomial \( \tilde{N}_1(z) \).

• Since \( \tilde{N}(z) \) is of degree at most \( d \) and \( A\tilde{D}(z) \) is of degree exactly \( d \), \( \tilde{N}_1 \) is a polynomial of degree \( d - 1 \). It now suffices to choose \( \tilde{P}(z) = \tilde{N}_1(z) \).

Now back to

\[
\frac{N(x)}{(x-a_1) \times \cdots \times (x-a_d)}
\]

Apply Lemma 1.10.13, with \( D(x) = (x-a_2) \times \cdots \times (x-a_d) \) and \( a = a_1 \). It says

\[
\frac{N(x)}{(x-a_1) \times \cdots \times (x-a_d)} = \frac{A_1}{x-a_1} + \frac{P(x)}{(x-a_2) \times \cdots \times (x-a_d)}
\]
for some polynomial $P$ of degree at most $d - 2$ and some number $A_1$.

Apply Lemma 1.10.13 a second time, with $D(x) = (x - a_3) \times \cdots \times (x - a_d)$, $N(x) = P(x)$ and $a = a_2$. It says
\[
\frac{P(x)}{(x - a_2) \times \cdots \times (x - a_d)} = \frac{A_2}{x - a_2} + \frac{Q(x)}{(x - a_3) \times \cdots \times (x - a_d)}
\]
for some polynomial $Q$ of degree at most $d - 3$ and some number $A_2$.

At this stage, we know that
\[
\frac{N(x)}{(x - a_1) \times \cdots \times (x - a_d)} = \frac{A_1}{x - a_1} + \frac{A_2}{x - a_2} + \frac{Q(x)}{(x - a_3) \times \cdots \times (x - a_d)}
\]
If we just keep going, repeatedly applying Lemma 1, we eventually end up with
\[
\frac{N(x)}{(x - a_1) \times \cdots \times (x - a_d)} = \frac{A_1}{x - a_1} + \cdots + \frac{A_d}{x - a_d}
\]
as required.

The general case with linear factors

Now consider splitting
\[
\frac{N(x)}{(x - a_1)^{n_1} \times \cdots \times (x - a_d)^{n_d}}
\]
into a sum of the form\(^{69}\)
\[
\left[ \frac{A_{1,1}}{x - a_1} \right] + \cdots + \left[ \frac{A_{1,n_1}}{(x - a_1)^{n_1}} \right] + \cdots + \left[ \frac{A_{d,1}}{x - a_d} \right] + \cdots + \left[ \frac{A_{d,n_d}}{(x - a_d)^{n_d}} \right]
\]
We now show that this decomposition can always be achieved, under the assumptions that the $a_i$'s are all different and $N(x)$ is a polynomial of degree at most $n_1 + \cdots + n_d - 1$. To do so, we shall repeatedly apply the following Lemma.

\textbf{Lemma 1.10.14.}

Let $N(x)$ and $D(x)$ be polynomials of degree $n$ and $d$ respectively, with $n < d + m$. Suppose that $a$ is NOT a zero of $D(x)$. Then there is a polynomial $P(x)$ of degree $p < d$ and numbers $A_1, \cdots, A_m$ such that
\[
\frac{N(x)}{D(x) (x - a)^m} = \frac{P(x)}{D(x)} + \frac{A_1}{x - a} + \frac{A_2}{(x - a)^2} + \cdots + \frac{A_m}{(x - a)^m}
\]

\(^{69}\) If we allow ourselves to use complex numbers as roots, this is the general case. We don’t need to consider quadratic (or higher) factors since all polynomials can be written as products of linear factors with complex coefficients.
Proof. As we did in the proof of the previous lemma, we write \( z = x - a \). Then \( \tilde{N}(z) = N(z + a) \) and \( \tilde{D}(z) = D(z + a) \) are polynomials of degree \( n \) and \( d \) respectively, \( \tilde{D}(0) = D(a) \neq 0 \).

- In order to complete the proof we have to find a polynomial \( \tilde{P}(z) \) of degree \( p < d \) and numbers \( A_1, \cdots, A_m \) such that

\[
\frac{\tilde{N}(z)}{\tilde{D}(z)z^m} = \frac{\tilde{P}(z)}{\tilde{D}(z)} + \frac{A_1}{z} + \frac{A_2}{z^2} + \cdots + \frac{A_m}{z^m}
\]

or equivalently, such that

\[
\tilde{P}(z)z^m + A_1z^{m-1}\tilde{D}(z) + A_2z^{m-2}\tilde{D}(z) + \cdots + A_m\tilde{D}(z) = \tilde{N}(z)
\]

- Now look at the polynomial on the left hand side. Every single term on the left hand side, except for the very last one, \( A_m\tilde{D}(z) \), has at least one power of \( z \). So the constant term on the left hand side is exactly the constant term in \( A_m\tilde{D}(z) \), which is equal to \( A_m\tilde{D}(0) \). The constant term on the right hand side is equal to \( \tilde{N}(0) \). So the constant terms on the left and right hand sides are the same if we choose \( A_m = \frac{\tilde{N}(0)}{\tilde{D}(0)} \). Recall that \( \tilde{D}(0) \neq 0 \) so \( A_m \) is well defined.

- Now move \( A_m\tilde{D}(z) \) to the right hand side.

\[
\tilde{P}(z)z^m + A_1z^{m-1}\tilde{D}(z) + A_2z^{m-2}\tilde{D}(z) + \cdots + A_{m-1}\tilde{D}(z) = \tilde{N}(z) - A_m\tilde{D}(z)
\]

The constant terms in \( \tilde{N}(z) \) and \( A_m\tilde{D}(z) \) are the same, so the right hand side contains no constant term and the right hand side is of the form \( \tilde{N}_1(z)z \) with \( \tilde{N}_1 \) a polynomial of degree at most \( d + m - 2 \). (Recall that \( \tilde{N} \) is of degree at most \( d + m - 1 \) and \( \tilde{D} \) is of degree at most \( d \).) Divide the whole equation by \( z \) to get

\[
\tilde{P}(z)z^{m-1} + A_1z^{m-2}\tilde{D}(z) + A_2z^{m-3}\tilde{D}(z) + \cdots + A_{m-1}\tilde{D}(z) = \tilde{N}_1(z).
\]

- Now, we can repeat the previous argument. The constant term on the left hand side, which is exactly equal to \( A_{m-1}\tilde{D}(0) \) matches the constant term on the right hand side, which is equal to \( \tilde{N}_1(0) \) if we choose \( A_{m-1} = \frac{\tilde{N}_1(0)}{\tilde{D}(0)} \). With this choice of \( A_{m-1} \)

\[
\tilde{P}(z)z^{m-1} + A_1z^{m-2}\tilde{D}(z) + A_2z^{m-3}\tilde{D}(z) + \cdots + A_{m-2}\tilde{D}(z)
\]

with \( \tilde{N}_2 \) a polynomial of degree at most \( d + m - 3 \). Divide by \( z \) and continue.

- After \( m \) steps like this, we end up with

\[
\tilde{P}(z)z = \tilde{N}_{m-1}(z) - A_1\tilde{D}(z)
\]

after having chosen \( A_1 = \frac{\tilde{N}_{m-1}(0)}{\tilde{D}(0)} \).
There is no constant term on the right side so that \( \tilde{N}_{m-1}(z) - A_1D(z) \) is of the form \( \tilde{N}_m(z)z \) with \( \tilde{N}_m \) a polynomial of degree \( d - 1 \). Choosing \( \tilde{P}(z) = \tilde{N}_m(z) \) completes the proof.

Now back to

\[
\frac{N(x)}{(x - a_1)^{n_1} \times \cdots \times (x - a_d)^{n_d}}
\]

Apply Lemma 1.10.14, with \( D(x) = (x - a_2)^{n_2} \times \cdots \times (x - a_d)^{n_d}, m = n_1 \) and \( a = a_1 \). It says

\[
\frac{N(x)}{(x - a_1)^{n_1} \times \cdots \times (x - a_d)^{n_d}} = \frac{A_{1,1}}{x - a_1} + \frac{A_{1,2}}{(x - a_1)^2} + \cdots + \frac{A_{1,n_1}}{(x - a_1)^{n_1}} + \frac{P(x)}{(x - a_2)^{n_2} \times \cdots \times (x - a_d)^{n_d}}
\]

Apply Lemma 1.10.14 a second time, with \( D(x) = (x - a_3)^{n_3} \times \cdots \times (x - a_d)^{n_d}, N(x) = P(x), m = n_2 \) and \( a = a_2 \). And so on. Eventually, we end up with

\[
\left[ \frac{A_{1,1}}{x - a_1} + \cdots + \frac{A_{1,n_1}}{(x - a_1)^{n_1}} \right] + \cdots + \left[ \frac{A_{d,1}}{x - a_d} + \cdots + \frac{A_{d,n_d}}{(x - a_d)^{n_d}} \right]
\]

which is exactly what we were trying to show.

1.11 Numerical Integration

By now the reader will have come to appreciate that integration is generally quite a bit more difficult than differentiation. There are a great many simple-looking integrals, such as \( \int e^{-x^2} \, dx \), that are either very difficult or even impossible to express in terms of standard functions\(^{70}\). Such integrals are not merely mathematical curiosities, but arise very naturally in many contexts. For example, the error function

\[
erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \, dt
\]

is extremely important in many areas of mathematics, and also in many practical applications of statistics.

In such applications we need to be able to evaluate this integral (and many others) at a given numerical value of \( x \). In this section we turn to the problem of how to find (approximate) numerical values for integrals, without having to evaluate them algebraically. To develop these methods we return to Riemann sums and our geometric interpretation of the definite integral as the signed area.

We start by describing (and applying) three simple algorithms for generating, numerically, approximate values for the definite integral \( \int_a^b f(x) \, dx \). In each algorithm, we begin in much the same way as we approached Riemann sums.

\(^{70}\) We apologise for being a little sloppy here — but we just want to say that it can be very hard or even impossible to write some integrals as some finite sized expression involving polynomials, exponentials, logarithms and trigonometric functions. We don’t want to get into a discussion of computability, though that is a very interesting topic.
• We first select an integer \( n > 0 \), called the “number of steps”.

• We then divide the interval of integration, \( a \leq x \leq b \), into \( n \) equal subintervals, each of length \( \Delta x = \frac{b-a}{n} \). The first subinterval runs from \( x_0 = a \) to \( x_1 = a + \Delta x \). The second runs from \( x_1 \) to \( x_2 = a + 2\Delta x \), and so on. The last runs from \( x_{n-1} = b - \Delta x \) to \( x_n = b \).

This splits the original integral into \( n \) pieces:

\[
\int_a^b f(x) \, dx = \int_{x_0}^{x_1} f(x) \, dx + \int_{x_1}^{x_2} f(x) \, dx + \cdots + \int_{x_{n-1}}^{x_n} f(x) \, dx
\]

Each subintegral \( \int_{x_{j-1}}^{x_j} f(x) \, dx \) is approximated by the area of a simple geometric figure. The three algorithms we consider approximate the area by rectangles, trapezoids and parabolas (respectively).

We will explain these rules in detail below, but we give a brief overview here:
The midpoint rule approximates each subintegral by the area of a rectangle of height

\[ \int_{x_{j-1}}^{x_j} f(x) \, dx \approx f \left( \frac{x_{j-1} + x_j}{2} \right) \Delta x \]

This is illustrated in the leftmost figure above.

The trapezoidal rule approximates each subintegral by the area of a trapezoid with

vertices at \((x_{j-1}, 0), (x_{j-1}, f(x_{j-1})), (x_j, f(x_j)), (x_j, 0)\):

\[ \int_{x_{j-1}}^{x_j} f(x) \, dx \approx \frac{1}{2} (f(x_{j-1}) + f(x_j)) \Delta x \]

The trapezoid is illustrated in the middle figure above. We shall derive the formula for the area shortly.

Simpson’s rule approximates two adjacent subintegrals by the area under a parabola that passes through the points \((x_{j-1}, f(x_{j-1})), (x_j, f(x_j))\) and \((x_{j+1}, f(x_{j+1})):\)

\[ \int_{x_{j-1}}^{x_{j+1}} f(x) \, dx \approx \frac{1}{3} (f(x_{j-1}) + 4f(x_j) + f(x_{j+1})) \Delta x \]

The parabola is illustrated in the right hand figure above. We shall derive the formula for the area shortly.

In what follows we need to refer to the midpoint between \(x_{j-1}\) and \(x_j\) very frequently. To save on writing (and typing) we introduce the notation

\[ \bar{x}_j = \frac{1}{2} (x_{j-1} + x_j). \]

**Notation 1.11.1 (Midpoints).**

1.11.1 The midpoint rule

The integral \( \int_{x_{j-1}}^{x_j} f(x) \, dx \) represents the area between the curve \( y = f(x) \) and the \( x\)-axis with \( x \) running from \( x_{j-1} \) to \( x_j \). The width of this region is \( x_j - x_{j-1} = \Delta x \). The height varies over the different values that \( f(x) \) takes as \( x \) runs from \( x_{j-1} \) to \( x_j \).

The midpoint rule approximates this area by the area of a rectangle of width \( x_j - x_{j-1} = \Delta x \) and height \( f(\bar{x}_j) \) which is the exact height at the midpoint of the range covered by \( x \).
The area of the approximating rectangle is \( f(\bar{x}_j)\Delta x \), and the midpoint rule approximates each subintegral by

\[
\int_{x_{j-1}}^{x_j} f(x) \, dx \approx f(\bar{x}_j)\Delta x
\]

Applying this approximation to each subinterval and summing gives us the following approximation of the full integral:

\[
\int_a^b f(x) \, dx = \int_{x_0}^{x_1} f(x) \, dx + \int_{x_1}^{x_2} f(x) \, dx + \cdots + \int_{x_{n-1}}^{x_n} f(x) \, dx
\]

\[
\approx f(\bar{x}_1)\Delta x + f(\bar{x}_2)\Delta x + \cdots + f(\bar{x}_n)\Delta x
\]

So notice that the approximation is the sum of the function evaluated at the midpoint of each interval and then multiplied by \( \Delta x \). Our other approximations will have similar forms.

In summary:

The midpoint rule approximation is

\[
\int_a^b f(x) \, dx \approx \left[ f(\bar{x}_1) + f(\bar{x}_2) + \cdots + f(\bar{x}_n) \right] \Delta x
\]

where \( \Delta x = \frac{b-a}{n} \) and

\[
x_0 = a \quad x_1 = a + \Delta x \quad x_2 = a + 2\Delta x \quad \cdots \quad x_{n-1} = b - \Delta x \quad x_n = b
\]

\[
\bar{x}_1 = \frac{x_0 + x_1}{2} \quad \bar{x}_2 = \frac{x_1 + x_2}{2} \quad \cdots \quad \bar{x}_{n-1} = \frac{x_{n-2} + x_{n-1}}{2} \quad \bar{x}_n = \frac{x_{n-1} + x_n}{2}
\]

Example 1.11.3 \( \int_0^1 \frac{4}{1+x^2} \, dx \)

We approximate the above integral using the midpoint rule with \( n = 8 \) step.

Solution.

- First we set up all the \( x \)-values that we will need. Note that \( a = 0, b = 1, \Delta x = \frac{1}{8} \) and

\[
x_0 = 0 \quad x_1 = \frac{1}{8} \quad x_2 = \frac{2}{8} \quad \cdots \quad x_7 = \frac{7}{8} \quad x_8 = \frac{8}{8} = 1
\]
Consequently
\[ \bar{x}_1 = \frac{1}{16} \quad \bar{x}_2 = \frac{3}{16} \quad \bar{x}_3 = \frac{5}{16} \quad \cdots \quad \bar{x}_8 = \frac{15}{16} \]

- We now apply Equation (1.11.2) to the integrand \( f(x) = \frac{4}{1 + x^2} \):

\[
\int_0^1 \frac{4}{1 + x^2} \, dx \approx \left[ \frac{f(\bar{x}_1)}{4} + \frac{f(\bar{x}_2)}{4} + \cdots + \frac{f(\bar{x}_n)}{4} \right] \Delta x
\]

\[
= \left[ \frac{4}{1 + \frac{1}{16^2}} + \frac{4}{1 + \frac{3^2}{16^2}} + \frac{4}{1 + \frac{5^2}{16^2}} + \cdots + \frac{4}{1 + \frac{13^2}{16^2}} + \frac{4}{1 + \frac{15^2}{16^2}} \right] \frac{1}{8}
\]

\[
= [3.98444 + 3.86415 + 3.64413 + 3.35738 + 3.03858 + 2.71618 + 2.40941 + 2.12890] \frac{1}{8}
\]

\[
= 3.1429
\]

where we have rounded to four decimal places.

- In this case we can compute the integral exactly (which is one of the reasons it was chosen as a first example):

\[
\int_0^1 \frac{4}{1 + x^2} \, dx = 4 \arctan x \bigg|_0^1 = \pi
\]

- So the error in the approximation generated by eight steps of the midpoint rule is

\[
|3.1429 - \pi| = 0.0013
\]

- The relative error is then

\[
\frac{|\text{approximate} - \text{exact}|}{\text{exact}} = \frac{|3.1429 - \pi|}{\pi} = 0.0004
\]

That is the error is 0.0004 times the actual value of the integral.

- We can write this as a percentage error by multiplying it by 100

\[
\text{percentage error} = 100 \times \frac{|\text{approximate} - \text{exact}|}{\text{exact}} = 0.04\%
\]

That is, the error is about 0.04% of the exact value.

Example 1.11.3

The midpoint rule gives us quite good estimates of the integral without too much work — though it is perhaps a little tedious to do by hand. Of course, it would be very helpful to quantify what we mean by “good” in this context and that requires us to discuss errors.

71 Thankfully it is very easy to write a program to apply the midpoint rule.
Suppose that \( \alpha \) is an approximation to \( A \). This approximation has

- absolute error \( |A - \alpha| \)
- relative error \( \frac{|A - \alpha|}{A} \)
- percentage error \( 100 \frac{|A - \alpha|}{A} \)

We will discuss errors further in Section 1.11.4 below.

### Example 1.11.5 (\( \int_0^\pi \sin x \, dx \))

As a second example, we apply the midpoint rule with \( n = 8 \) steps to the above integral.

- We again start by setting up all the \( x \)-values that we will need. So \( a = 0 \), \( b = \pi \), \( \Delta x = \frac{\pi}{8} \) and

\[
\begin{align*}
x_0 &= 0, & x_1 &= \frac{\pi}{8}, & x_2 &= \frac{2\pi}{8}, & \cdots & x_7 &= \frac{7\pi}{8}, & x_8 &= \frac{8\pi}{8} = \pi
\end{align*}
\]

Consequently,

\[
\begin{align*}
x_1 &= \frac{\pi}{16}, & x_2 &= \frac{3\pi}{16}, & \cdots & x_7 &= \frac{13\pi}{16}, & x_8 &= \frac{15\pi}{16}
\end{align*}
\]

- Now apply Equation (1.11.2) to the integrand \( f(x) = \sin x \):

\[
\int_0^\pi \sin x \, dx \approx \left[ \sin(x_1) + \sin(x_2) + \cdots + \sin(x_8) \right] \Delta x
\]

\[
= \left[ \sin\left(\frac{\pi}{16}\right) + \sin\left(\frac{3\pi}{16}\right) + \sin\left(\frac{5\pi}{16}\right) + \sin\left(\frac{7\pi}{16}\right) + \sin\left(\frac{9\pi}{16}\right) + \sin\left(\frac{11\pi}{16}\right) + \sin\left(\frac{13\pi}{16}\right) + \sin\left(\frac{15\pi}{16}\right) \right] \frac{\pi}{8}
\]

\[
= \left[ 0.1951 + 0.5556 + 0.8315 + 0.9808 + 0.9808 + 0.8315 + 0.5556 + 0.1951 \right] \times 0.3927
\]

\[
= 5.1260 \times 0.3927 = 2.013
\]

- Again, we have chosen this example so that we can compare it against the exact value:

\[
\int_0^\pi \sin x \, dx = \left[ -\cos x \right]_0^\pi = -\cos \pi + \cos 0 = 2.
\]

- So with eight steps of the midpoint rule we achieved

\[
\begin{align*}
\text{absolute error} &= |2.013 - 2| = 0.013 \\
\text{relative error} &= \frac{|2.013 - 2|}{2} = 0.0065 \\
\text{percentage error} &= 100 \times \frac{|2.013 - 2|}{2} = 0.65\%
\end{align*}
\]

With little work we have managed to estimate the integral to within 1% of its true value.
1.11.2 The trapezoidal rule

Consider again the area represented by the integral \( \int_{x_{j-1}}^{x_j} f(x) \, dx \). The trapezoidal rule \(^{72}\) (unsurprisingly) approximates this area by a trapezoid \(^{73}\) whose vertices lie at 
\((x_{j-1}, 0), (x_{j-1}, f(x_{j-1})), (x_j, f(x_j))\) and \((x_j, 0)\).

The trapezoidal approximation of the integral \( \int_{x_{j-1}}^{x_j} f(x) \, dx \) is the shaded region in the figure on the right above. It has width \( x_j - x_{j-1} = \Delta x \). Its left hand side has height \( f(x_{j-1}) \) and its right hand side has height \( f(x_j) \).

As the figure below shows, the area of a trapezoid is its width times its average height.

So the trapezoidal rule approximates each subintegral by

\[
\int_{x_{j-1}}^{x_j} f(x) \, dx \approx \frac{f(x_{j-1}) + f(x_j)}{2} \Delta x
\]

Applying this approximation to each subinterval and then summing the result gives us the following approximation of the full integral

\[
\int_a^b f(x) \, dx = \int_{x_0}^{x_1} f(x) \, dx + \int_{x_1}^{x_2} f(x) \, dx + \cdots + \int_{x_{n-1}}^{x_n} f(x) \, dx
\approx \frac{f(x_0) + f(x_1)}{2} \Delta x + \frac{f(x_1) + f(x_2)}{2} \Delta x + \cdots + \frac{f(x_{n-1}) + f(x_n)}{2} \Delta x
= \left[ \frac{1}{2} f(x_0) + f(x_1) + f(x_2) + \cdots + f(x_{n-1}) + \frac{1}{2} f(x_n) \right] \Delta x
\]

\(^{72}\) This method is also called the “trapezoid rule” and “trapezium rule”.

\(^{73}\) A trapezoid is a four sided polygon, like a rectangle. But, unlike a rectangle, the top and bottom of a trapezoid need not be parallel.
So notice that the approximation has a very similar form to the midpoint rule, excepting that

- we evaluate the function at the \( x_j \)’s rather than at the midpoints, and
- we multiply the value of the function at the endpoints \( x_0, x_n \) by \( \frac{1}{2} \).

In summary:

**Equation 1.11.6 (The trapezoidal rule).**

The trapezoidal rule approximation is

\[
\int_a^b f(x) \, dx \approx \left[ \frac{1}{2} f(x_0) + f(x_1) + f(x_2) + \cdots + f(x_{n-1}) + \frac{1}{2} f(x_n) \right] \Delta x
\]

where

\[
\Delta x = \frac{b-a}{n}, \quad x_0 = a, \quad x_1 = a + \Delta x, \quad x_2 = a + 2\Delta x, \quad \ldots, \quad x_{n-1} = b - \Delta x, \quad x_n = b
\]

To compare and contrast we apply the trapezoidal rule to the examples we did above with the midpoint rule.

**Example 1.11.7** \( \int_0^1 \frac{4}{1+x^2} \, dx \) — using the trapezoidal rule

**Solution.** We proceed very similarly to Example 1.11.3 and again use \( n = 8 \) steps.

- We again have \( f(x) = \frac{4}{1+x^2}, a = 0, b = 1, \Delta x = \frac{1}{8} \) and

  \[
  x_0 = 0 \quad x_1 = \frac{1}{8} \quad x_2 = \frac{2}{8} \quad \cdots \quad x_7 = \frac{7}{8} \quad x_8 = \frac{8}{8} = 1
  \]

- Applying the trapezoidal rule, Equation (1.11.6), gives

\[
\int_0^1 \frac{4}{1+x^2} \, dx \approx \left[ \frac{1}{2} \frac{4}{1 + x_0^2} + \frac{4}{1 + x_1^2} + \cdots + \frac{4}{1 + x_7^2} + \frac{1}{2} \frac{4}{1 + x_8^2} \right] \Delta x
\]

\[
= \left[ \frac{1}{2} \frac{4}{1 + 0^2} + \frac{4}{1 + \frac{1}{8}^2} + \frac{4}{1 + \frac{2}{8}^2} + \frac{4}{1 + \frac{3}{8}^2} + \frac{4}{1 + \frac{4}{8}^2} + \frac{4}{1 + \frac{5}{8}^2} + \frac{4}{1 + \frac{6}{8}^2} + \frac{4}{1 + \frac{7}{8}^2} + \frac{1}{2} \frac{4}{1 + 1^2} \right] \frac{1}{8}
\]

\[
= \left[ \frac{1}{2} \times 4 + 3.939 + 3.765 + 3.507 + 3.2 + 2.876 + 2.56 + 2.266 + \frac{1}{2} \times 2 \right] \frac{1}{8}
\]

\[
= 3.139
\]

to three decimal places.
The exact value of the integral is still $\pi$. So the error in the approximation generated by eight steps of the trapezoidal rule is $|3.139 - \pi| = 0.0026$, which is $100 \frac{|3.139 - \pi|}{\pi} \% = 0.08\%$ of the exact answer. Notice that this is roughly twice the error that we achieved using the midpoint rule in Example 1.11.3.

Example 1.11.7

Let us also redo Example 1.11.5 using the trapezoidal rule.

Example 1.11.8 ([\int_0^\pi \sin x \, dx] — using the trapezoidal rule)

Solution. We proceed very similarly to Example 1.11.5 and again use $n = 8$ steps.

• We again have $a = 0$, $b = \pi$, $\Delta x = \frac{\pi}{8}$ and
  
  \begin{align*}
  x_0 &= 0 \\
  x_1 &= \frac{\pi}{8} \\
  x_2 &= \frac{2\pi}{8} \\
  x_3 &= \frac{3\pi}{8} \\
  & \vdots \\
  x_7 &= \frac{7\pi}{8} \\
  x_8 &= \frac{8\pi}{8} = \pi
  \end{align*}

• Applying the trapezoidal rule, Equation (1.11.6), gives

\[
\int_0^\pi \sin x \, dx \approx \left[ \frac{1}{2} \sin(x_0) + \sin(x_1) + \cdots + \sin(x_7) + \frac{1}{2} \sin(x_8) \right] \Delta x \\
= \left[ \frac{1}{2} \sin 0 + \sin \frac{\pi}{8} + \sin \frac{2\pi}{8} + \sin \frac{3\pi}{8} + \sin \frac{4\pi}{8} + \sin \frac{5\pi}{8} + \sin \frac{6\pi}{8} + \sin \frac{7\pi}{8} + \frac{1}{2} \sin \frac{8\pi}{8} \right] \frac{\pi}{8} \\
= \left[ \frac{1}{2} \times 0 + 0.3827 + 0.7071 + 0.9239 + 1.0000 + 0.9239 + 0.7071 + 0.3827 + \frac{1}{2} \times 0 \right] \times 0.3927 \\
= 5.0274 \times 0.3927 = 1.974
\]

• The exact answer is $\int_0^\pi \sin x \, dx = -\cos x \bigg|_0^\pi = 2$. So with eight steps of the trapezoidal rule we achieved $100 \frac{|1.974 - 2|}{2} = 1.3\%$ accuracy. Again this is approximately twice the error we achieved in Example 1.11.5 using the midpoint rule.

Example 1.11.8

These two examples suggest that the midpoint rule is more accurate than the trapezoidal rule. Indeed, this observation is born out by a rigorous analysis of the error — see Section 1.11.4.

1.11.3 Simpson’s Rule

When we use the trapezoidal rule we approximate the area $\int_{x_{j-1}}^{x_j} f(x) \, dx$ by the area between the $x$-axis and a straight line that runs from $(x_{j-1}, f(x_{j-1}))$ to $(x_j, f(x_j))$ — that is, we approximate the function $f(x)$ on this interval by a linear function that agrees with the function at each endpoint. An obvious way to extend this — just as we did when extending linear approximations to quadratic approximations in our differential calculus course
— is to approximate the function with a quadratic. This is precisely what Simpson’s rule does.

Simpson’s rule approximates the integral over two neighbouring subintervals by the area between a parabola and the x-axis. In order to describe this parabola we need 3 distinct points (which is why we approximate two subintegrals at a time). That is, we approximate

\[
\int_{x_0}^{x_1} f(x) \, dx + \int_{x_1}^{x_2} f(x) \, dx = \int_{x_0}^{x_2} f(x) \, dx
\]

by the area bounded by the parabola that passes through the three points \((x_0, f(x_0))\), \((x_1, f(x_1))\) and \((x_2, f(x_2))\), the x-axis and the vertical lines \(x = x_0\) and \(x = x_2\). We repeat this on the next pair of subintervals and approximate \(\int_{x_2}^{x_4} f(x) \, dx\) by the area between the x–axis and the part of a parabola with \(x \leq x \leq x_4\). This parabola passes through the three points \((x_2, f(x_2))\), \((x_3, f(x_3))\) and \((x_4, f(x_4))\). And so on. Because Simpson’s rule does the approximation two slices at a time, \(n\) must be even.

To derive Simpson’s rule formula, we first find the equation of the parabola that passes through the three points \((x_0, f(x_0))\), \((x_1, f(x_1))\) and \((x_2, f(x_2))\). Then we find the area between the x–axis and the part of that parabola with \(x_0 \leq x \leq x_2\). To simplify this computation consider a parabola passing through the points \((-h, y_{-1})\), \((0, y_0)\) and \((h, y_1)\).

Write the equation of the parabola as

\[ y = Ax^2 + Bx + C \]

Then the area between it and the x-axis with \(x\) running from \(-h\) to \(h\) is

\[
\int_{-h}^{h} [Ax^2 + Bx + C] \, dx = \left[ \frac{A}{3} x^3 + \frac{B}{2} x^2 + Cx \right]_{-h}^{h}
\]

\[= \frac{2A}{3} h^3 + 2Ch \quad \text{it is helpful to write it as}
\]

\[= \frac{h}{3} \left( 2Ah^2 + 6C \right) \]

\[74\] Simpson’s rule is named after the 18th century English mathematician Thomas Simpson, despite its use a century earlier by the German mathematician and astronomer Johannes Kepler. In many German texts the rule is often called Kepler’s rule.
Now, the three points \((-h, y_{-1}), (0, y_0)\) and \((h, y_1)\) lie on this parabola if and only if
\[
\begin{align*}
Ah^2 - Bh + C &= y_{-1} & \text{at } (-h, y_{-1}) \\
C &= y_0 & \text{at } (0, y_0) \\
Ah^2 + Bh + C &= y_1 & \text{at } (h, y_1)
\end{align*}
\]
Adding the first and third equations together gives us
\[
2Ah^2 + (B - B)h + 2C = y_{-1} + y_1
\]
To this we add four times the middle equation
\[
2Ah^2 + 6C = y_{-1} + 4y_0 + y_1.
\]
This means that
\[
\text{area} = \int_{-h}^{h} [Ax^2 + Bx + C] \, dx = \frac{h}{3} \left( 2Ah^2 + 6C \right)
\]
\[
= \frac{h}{3} (y_{-1} + 4y_0 + y_1)
\]
Note that here

- \(h\) is one half of the length of the \(x\)-interval under consideration
- \(y_{-1}\) is the height of the parabola at the left hand end of the interval under consideration
- \(y_0\) is the height of the parabola at the middle point of the interval under consideration
- \(y_1\) is the height of the parabola at the right hand end of the interval under consideration

So Simpson’s rule approximates
\[
\int_{x_0}^{x_2} f(x) \, dx \approx \frac{1}{3} \Delta x [f(x_0) + 4f(x_1) + f(x_2)]
\]
and
\[
\int_{x_2}^{x_4} f(x) \, dx \approx \frac{1}{3} \Delta x [f(x_2) + 4f(x_3) + f(x_4)]
\]
and so on. Summing these all together gives:
\[
\int_{a}^{b} f(x) \, dx = \int_{x_0}^{x_2} f(x) \, dx + \int_{x_2}^{x_4} f(x) \, dx + \int_{x_4}^{x_6} f(x) \, dx + \cdots + \int_{x_{n-2}}^{x_n} f(x) \, dx
\]
\[
\approx \frac{\Delta x}{3} [f(x_0) + 4f(x_1) + f(x_2)] + \frac{\Delta x}{3} [f(x_2) + 4f(x_3) + f(x_4)]
\]
\[
+ \frac{\Delta x}{3} [f(x_4) + 4f(x_5) + f(x_6)] + \cdots + \frac{\Delta x}{3} [f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)]
\]
\[
= \left[ f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \cdots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n) \right] \frac{\Delta x}{3}
\]
In summary
The Simpson’s rule approximation is

\[ \int_a^b f(x) \, dx \approx \left[ f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \cdots \right. \\
\left. \cdots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n) \right] \frac{\Delta x}{3} \]

where \( n \) is even and \( \Delta x = \frac{b-a}{n} \), \( x_0 = a \), \( x_1 = a + \Delta x \), \( x_2 = a + 2\Delta x \), \( \cdots \), \( x_{n-1} = b - \Delta x \), \( x_n = b \).

Equation 1.11.9 (Simpson’s rule).

Notice that Simpson’s rule requires essentially no more work than the trapezoidal rule. In both rules we must evaluate \( f(x) \) at \( x = x_0, x_1, \ldots, x_n \), but we add those terms multiplied by different constants.

Let’s put it to work on our two running examples.

**Example 1.11.10** \( \int_0^1 \frac{4}{1+x^2} \, dx \) — using Simpson’s rule

**Solution.** We proceed almost identically to Example 1.11.7 and again use \( n = 8 \) steps.

- We have the same \( \Delta, a, b, x_0, \ldots, x_n \) as Example 1.11.7.

- Applying Equation 1.11.9 gives

\[ \int_0^1 \frac{4}{1+x^2} \, dx \approx \left[ \frac{4}{1+0^2} + 4\frac{4}{1+1/8} + 2\frac{4}{1+2/8} + 4\frac{4}{1+3^2/8} \\
+ 2\frac{4}{1+4^2/8} + \frac{4}{1+5^2/8} + 2\frac{4}{1+6^2/8} + 4\frac{4}{1+7^2/8} + \frac{4}{1+8^2/8} \right] \frac{1}{8 \times 3} \]

\[ = \left[ 4 + 4 \times 3.938461538 + 2 \times 3.764705882 + 4 \times 3.506849315 \\
+ 2 \times 3.2 + 4 \times 2.876404494 + 2 \times 2.56 + 4 \times 2.265486726 + 2 \right] \frac{1}{8 \times 3} \]

\[ = 3.14159250 \]

to eight decimal places.

- This agrees with \( \pi \) (the exact value of the integral) to six decimal places. So the error in the approximation generated by eight steps of Simpson’s rule is \( |3.14159250 - \pi| = 1.5 \times 10^{-7} \), which is \( 100 \frac{|3.14159250 - \pi|}{\pi} \% = 5 \times 10^{-6} \% \) of the exact answer.

75 There is an easy generalisation of Simpson’s rule that uses cubics instead of parabolas. It leads to the formula

\[ \int_a^b f(x) \, dx = \left[ f(x_0) + 3f(x_1) + 3f(x_2) + 2f(x_3) + 2f(x_4) + 3f(x_5) + 3f(x_6) + 2f(x_7) + \cdots + f(x_n) \right] \frac{3\Delta x}{8} \]

where \( n \) is a multiple of 3. This result is known as Simpson’s second rule and Simpson’s \( 3/8 \) rule. While one can push this approach further (using quartics, quintics etc), it can sometimes lead to larger errors — the interested reader should look up Runge’s phenomenon.
It is striking that the absolute error approximating with Simpson’s rule is so much smaller than the error from the midpoint and trapezoidal rules.

midpoint error = 0.0013  
trapezoid error = 0.0026  
Simpson error = 0.00000015

Buoyed by this success, we will also redo Example 1.11.8 using Simpson’s rule.

Example 1.11.11 \( \int_0^\pi \sin x \, dx \) — Simpson’s rule

Solution. We proceed almost identically to Example 1.11.8 and again use \( n = 8 \) steps.

- We have the same \( \Delta, a, b, x_0, \ldots, x_n \) as Example 1.11.7.

- Applying Equation 1.11.9 gives

\[
\int_0^\pi \sin x \, dx \approx \left[ \sin(x_0) + 4 \sin(x_1) + 2 \sin(x_2) + \cdots + 4 \sin(x_7) + \sin(x_8) \right] \frac{4\pi}{8x3}
\]

\[
= \left[ \sin(0) + 4 \sin\left(\frac{\pi}{8}\right) + 2 \sin\left(\frac{2\pi}{8}\right) + 4 \sin\left(\frac{3\pi}{8}\right) + 2 \sin\left(\frac{4\pi}{8}\right) \\
+ 4 \sin\left(\frac{5\pi}{8}\right) + 2 \sin\left(\frac{6\pi}{8}\right) + 4 \sin\left(\frac{7\pi}{8}\right) + \sin\left(\frac{8\pi}{8}\right) \right] \frac{\pi}{8x3}
\]

\[
= \left[ 0 + 4 \times 0.382683 + 2 \times 0.707107 + 4 \times 0.923880 + 2 \times 1.0 \\
+ 4 \times 0.923880 + 2 \times 0.707107 + 4 \times 0.382683 + 0 \right] \frac{\pi}{8x3}
\]

\[
= 15.280932 \times 0.130900
\]

\[
= 2.00027
\]

- With only eight steps of Simpson’s rule we achieved \( \frac{2.00027 - 2}{2} = 0.014\% \) accuracy.

Again we contrast the error we achieved with the other two rules:

midpoint error = 0.013  
trapezoid error = 0.026  
Simpson error = 0.00027

This completes our derivation of the midpoint, trapezoidal and Simpson’s rules for approximating the values of definite integrals. So far we have not attempted to see how efficient and how accurate the algorithms are in general. That’s our next task.
1.11.4 Three Simple Numerical Integrators – Error Behaviour

Now we are armed with our three (relatively simple) method for numerical integration we should give thought to how practical they might be in the real world\(^{76}\). Two obvious considerations when deciding whether or not a given algorithm is of any practical value are

(a) the amount of computational effort required to execute the algorithm and

(b) the accuracy that this computational effort yields.

For algorithms like our simple integrators, the bulk of the computational effort usually goes into evaluating the function \(f(x)\). The number of evaluations of \(f(x)\) required for \(n\) steps of the midpoint rule is \(n\), while the number required for \(n\) steps of the trapezoidal and Simpson’s rules is \(n + 1\). So all three of our rules require essentially the same amount of effort – one evaluation of \(f(x)\) per step.

To get a first impression of the error behaviour of these methods, we apply them to a problem whose answer we know exactly:

\[
\int_0^\pi \sin x \, dx = -\cos x\bigg|_0^\pi = 2.
\]

To be a little more precise, we would like to understand how the errors of the three methods change as we increase the effort we put in (as measured by the number of steps \(n\)). The following table lists the error in the approximate value for this number generated by our three rules applied with three different choices of \(n\). It also lists the number of evaluations of \(f\) required to compute the approximation.

<table>
<thead>
<tr>
<th>(n)</th>
<th>Midpoint</th>
<th>Trapezoidal</th>
<th>Simpson’s</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>error</td>
<td># evals</td>
<td>error</td>
</tr>
<tr>
<td>10</td>
<td>(4.1 \times 10^{-1})</td>
<td>10</td>
<td>(8.2 \times 10^{-1})</td>
</tr>
<tr>
<td>100</td>
<td>(4.1 \times 10^{-3})</td>
<td>100</td>
<td>(8.2 \times 10^{-3})</td>
</tr>
<tr>
<td>1000</td>
<td>(4.1 \times 10^{-5})</td>
<td>1000</td>
<td>(8.2 \times 10^{-5})</td>
</tr>
</tbody>
</table>

Observe that

- Using 101 evaluations of \(f\) worth of Simpson’s rule gives an error 80 times smaller than 1000 evaluations of \(f\) worth of the midpoint rule.

- The trapezoidal rule error with \(n\) steps is about twice the midpoint rule error with \(n\) steps.

- With the midpoint rule, increasing the number of steps by a factor of 10 appears to reduce the error by about a factor of \(10^2 = n^2\).

- With the trapezoidal rule, increasing the number of steps by a factor of 10 appears to reduce the error by about a factor of \(10^2 = n^2\).

\(^{76}\) Indeed, even beyond the “real world” of many applications in first year calculus texts, some of the methods we have described are used by actual people (such as ship builders, engineers and surveyors) to estimate areas and volumes of actual objects!
• With Simpson’s rule, increasing the number of steps by a factor of 10 appears to reduce the error by about a factor of $10^4 = n^4$.

So it looks like

\[
\begin{align*}
\text{approx value of } \int_a^b f(x) \, dx \text{ given by } n \text{ midpoint steps} & \approx \int_a^b f(x) \, dx + K_M \cdot \frac{1}{n^2} \\
\text{approx value of } \int_a^b f(x) \, dx \text{ given by } n \text{ trapezoidal steps} & \approx \int_a^b f(x) \, dx + K_T \cdot \frac{1}{n^2} \\
\text{approx value of } \int_a^b f(x) \, dx \text{ given by } n \text{ Simpson’s steps} & \approx \int_a^b f(x) \, dx + K_M \cdot \frac{1}{n^4}
\end{align*}
\]

with some constants $K_M$, $K_T$ and $K_S$. It also seems that $K_T \approx 2K_M$.

Figure 1.11.1.

A log-log plot of the error in the $n$ step approximation to $\int_0^\pi \sin x \, dx$. 

179
To test these conjectures for the behaviour of the errors we apply our three rules with about ten different choices of \( n \) of the form \( n = 2^m \) with \( m \) integer. Figure 1.11.1 contains two graphs of the results. The left-hand plot shows the results for the midpoint and trapezoidal rules and the right-hand plot shows the results for Simpson’s rule.

For each rule we are expecting (based on our conjectures above) that the error

\[ e_n = |\text{exact value} - \text{approximate value}| \]

with \( n \) steps is (roughly) of the form

\[ e_n = K \frac{1}{n^k} \]

for some constants \( K \) and \( k \). We would like to test if this is really the case, by graphing \( Y = e_n \) against \( X = n \) and seeing if the graph “looks right”. But it is not easy to tell whether or not a given curve really is \( Y = \frac{K}{X^k} \), for some specific \( k \), by just looking at it. However, your eye is pretty good at determining whether or not a graph is a straight line.

Fortunately, there is a little trick that turns the curve \( Y = \frac{K}{X^k} \) into a straight line – no matter what \( k \) is. Instead of plotting \( Y \) against \( X \), we plot \( \log Y \) against \( \log X \). This transformation works because when \( Y = \frac{K}{X^k} \)

\[ \log Y = \log K - k \log X \]

So plotting \( y = \log Y \) against \( x = \log X \) gives the straight line \( y = \log K - kx \), which has slope \(-k\) and \( y\)-intercept \( \log K \).

The three graphs in Figure 1.11.1 plot \( y = \log_2 e_n \) against \( x = \log_2 n \) for our three rules. Note that we have chosen to use logarithms with this “unusual base” because it makes it very clear how much the error is improved if we double the number of steps used. To be more precise — one unit step along the \( x\)-axis represents changing \( n \to 2n \). For example, applying Simpson’s rule with \( n = 2^4 \) steps results in an error of 0.000166, so the point \((x = \log_2 2^4 = 4, y = \log_2 0.000166 = \frac{\log 0.000166}{\log 2} = -15.8)\) has been included on the graph. Doubling the effort used — that is, doubling the number of steps to \( n = 2^5 \) — results in an error of 0.00000103. So, the data point \((x = \log_2 2^5 = 5, y = \log_2 0.00000103 = \frac{\ln 0.00000103}{\ln 2} = -19.9)\) lies on the graph. Note that the \( x\)-coordinates of these points differ by 1 unit.

77 There is a variant of this trick that works even when you don’t know the answer to the integral ahead of time. Suppose that you suspect that the approximation satisfies

\[ M_n = A + K \frac{1}{n^k} \]

where \( A \) is the exact value of the integral and suppose that you don’t know the values of \( A, K \) and \( k \). Then

\[ M_n - M_{2n} = K \frac{1}{n^k} - K \frac{1}{(2n)^k} = K \left(1 - \frac{1}{2^k}\right) \frac{1}{n^k} \]

so plotting \( y = \log(M_n - M_{2n}) \) against \( x = \log n \) gives the straight line \( y = \log \left[K \left(1 - \frac{1}{2^k}\right)\right] - kx \).

78 Now is a good time for a quick revision of logarithms — see “Whirlwind review of logarithms” in Section 2.7 of the CLP Mathematics 100 notes.
For each of the three sets of data points, a straight line has also been plotted “through” the data points. A procedure called linear regression has been used to decide precisely which straight line to plot. It provides a formula for the slope and $y$–intercept of the straight line which “best fits” any given set of data points. From the three lines, it sure looks like $k = 2$ for the midpoint and trapezoidal rules and $k = 4$ for Simpson’s rule. It also looks like the ratio between the value of $K$ for the trapezoidal rule, namely $K = 2^{0.7253}$, and the value of $K$ for the midpoint rule, namely $K = 2^{-0.2706}$, is pretty close to 2: $2^{0.7253} / 2^{-0.2706} = 2^{0.9959}$.

The intuition, about the error behaviour, that we have just developed is in fact correct — provided the integrand $f(x)$ is reasonably smooth. To be more precise

**Theorem 1.11.12 (Numerical integration errors).**

Assume that $|f''(x)| \leq M$ for all $a \leq x \leq b$. Then

the total error introduced by the midpoint rule is bounded by

$$\frac{M (b-a)^3}{24 n^2}$$

and

the total error introduced by the trapezoidal rule is bounded by

$$\frac{M (b-a)^3}{12 n^2}$$

when approximating $\int_a^b f(x) \, dx$. Further, if $|f^{(4)}(x)| \leq L$ for all $a \leq x \leq b$, then

the total error introduced by Simpson’s rule is bounded by

$$\frac{L (b-a)^5}{180 n^4}.$$ 

The first of these error bounds is proven in the following (optional) section. Here are some examples which illustrate how they are used. First let us check that the above result is consistent with our data in Figure 1.11.1

**Example 1.11.13 (Midpoint rule error approximating $\int_0^\pi \sin x \, dx$)**

- The integral $\int_0^\pi \sin x \, dx$ has $b - a = \pi$.
- The second derivative of the integrand satisfies

$$\left| \frac{d^2}{dx^2} \sin x \right| = |-\sin x| \leq 1$$

So we take $M = 1$. 

---

79 Linear regression is not part of this course as its derivation requires some multivariable calculus. It is a very standard technique in statistics.
• So the error, $e_n$, introduced when $n$ steps are used is bounded by

$$|e_n| \leq \frac{M(b-a)^3}{24 \frac{n^2}{24 n^2}} = \frac{\pi^3}{24 n^2} \approx 1.29 \frac{1}{n^2}$$

• The data in the graph in Figure 1.11.1 gives

$$|e_n| \approx 2^{-0.2706} \frac{1}{n^2} = 0.83 \frac{1}{n^2}$$

which is consistent with the bound $|e_n| \leq \frac{\pi^3}{24 n^2}$.

Example 1.11.13

In a typical application we would be asked to evaluate a given integral to some specified accuracy. For example, if you are manufacturer and your machinery can only cut materials to an accuracy of $\frac{1}{10}$ of a millimeter, there is no point in making design specifications more accurate than $\frac{1}{10}$ of a millimeter.

Example 1.11.14

Suppose, for example, that we wish to use the midpoint rule to evaluate

$$\int_0^1 e^{-x^2} \, dx$$

to within an accuracy of $10^{-6}$.

Solution.

• The integral has $a = 0$ and $b = 1$.

• The first two derivatives of the integrand are

$$\frac{d}{dx} e^{-x^2} = -2xe^{-x^2} \quad \text{and} \quad \frac{d^2}{dx^2} e^{-x^2} = \frac{d}{dx} (-2xe^{-x^2}) = -2e^{-x^2} + 4x^2 e^{-x^2} = 2(2x^2 - 1)e^{-x^2}$$

• As $x$ runs from 0 to 1, $2x^2 - 1$ increases from $-1$ to 1, so that

$$0 \leq x \leq 1 \implies |2x^2 - 1| \leq 1, \quad e^{-x^2} \leq 1 \implies |2(2x^2 - 1)e^{-x^2}| \leq 2$$

So we take $M = 2$. 80

---

80 This is our favourite running example of an integral that cannot be evaluated algebraically — we need to use numerical methods.
The error introduced by the \( n \) step midpoint rule is at most
\[
e_n \leq \frac{M (b-a)^3}{24 \frac{n^2}{n^2}} \leq \frac{2 (1-0)^3}{24 \frac{n^2}{n^2}} = \frac{1}{12n^2}
\]

We need this error to be smaller than \( 10^{-6} \) so
\[
e_n \leq \frac{1}{12n^2} \leq 10^{-6}
\]
and so
\[
12n^2 \geq 10^6
\]
clean up
\[
n^2 \geq \frac{10^6}{12} = 83333.3 \ldots
\]
square root both sides
\[
n \geq 288.7
\]
So 289 steps of the midpoint rule will do the job.

In fact \( n = 289 \) results in an error of about \( 3.7 \times 10^{-7} \).

Example 1.11.14
That seems like far too much work, and the trapezoidal rule will have twice the error. So we should look at Simpson's rule.

Example 1.11.15
Suppose now that we wish evaluate \( \int_0^1 e^{-x^2} \, dx \) to within an accuracy of \( 10^{-6} \) — but now using Simpson's rule. How many steps should we use?

Solution.

Again we have \( a = 0, b = 1 \).

We then need to bound \( \frac{d^4}{dx^4} e^{-x^2} \) on the domain of integration, \( 0 \leq x \leq 1 \).

\[
\frac{d^3}{dx^3} e^{-x^2} = \frac{d}{dx} \{2(2x^2 - 1)e^{-x^2}\} = 8xe^{-x^2} - 4x(2x^2 - 1)e^{-x^2}
\]
\[
= 4(-2x^3 + 3x)e^{-x^2}
\]
\[
\frac{d^4}{dx^4} e^{-x^2} = \frac{d}{dx} \{4(-2x^3 + 3x)e^{-x^2}\} = 4(-6x^2 + 3)e^{-x^2} - 8x(-2x^3 + 3x)e^{-x^2}
\]
\[
= 4(4x^4 - 12x^2 + 3)e^{-x^2}
\]

Now, for any \( x, e^{-x^2} \leq 1 \). Also, for \( 0 \leq x \leq 1 \),
\[
0 \leq x^2, x^4 \leq 1
\]
so
\[
3 \leq 4x^4 + 3 \leq 7
\]
and
\[
-12 \leq -12x^2 \leq 0
\]
adding these together gives
\[
-9 \leq 4x^4 - 12x^2 + 3 \leq 7
\]
Consequently, \(|4x^4 - 12x^2 + 3|\) is bounded by 9 and so

\[
\left| \frac{d^4}{dx^4} e^{-x^2} \right| \leq 4 \times 9 = 36
\]

So take \(L = 36\).

- The error introduced by the \(n\) step Simpson’s rule is at most

\[
e_n \leq \frac{L}{180} \frac{(b-a)^5}{n^4}
\]

\[
\leq \frac{36}{180} \frac{(1-0)^5}{n^4} = \frac{1}{5n^4}
\]

- In order for this error to be no more than \(10^{-6}\) we require \(n\) to satisfy

\[
e_n \leq \frac{1}{5n^4} \leq 10^{-6}
\]

\[5n^4 \geq 10^6\]

\[n^4 \geq 200000\quad \text{take fourth root}\]

\[n \geq 21.15\]

So 22 steps of Simpson’s rule will do the job.

- \(n = 22\) steps actually results in an error of \(3.5 \times 10^{-8}\). The reason that we get an error so much smaller than we need is that we have overestimated the number of steps required. This, in turn, occurred because we made quite a rough bound of \(\left| \frac{d^4}{dx^4} f(x) \right| \leq 36\). If we are more careful then we will get a slightly smaller \(n\). It actually turns out\(^{81}\) that you only need \(n = 10\) to approximate within \(10^{-6}\).

---

**Example 1.11.15**

**1.11.5 Optional — An error bound for the midpoint rule**

We now try develop some understanding as to why we got the above experimental results. We start with the error generated by a single step of the midpoint rule. That is, the error introduced by the approximation

\[
\int_{x_0}^{x_1} f(x) \, dx \approx f(\bar{x}_1) \Delta x \quad \text{where } \Delta x = x_1 - x_0, \ \bar{x}_1 = \frac{x_0 + x_1}{2}
\]

To do this we are going to need to apply integration by parts in a sneaky way. Let us start by considering\(^{82}\) a subinterval \(a \leq x \leq \beta\) and let’s call the width of the subinterval \(2q\) so

\[\]

---

81 The authors tested this empirically.  
82 We chose this interval so that we didn’t have lots of subscripts floating around in the algebra.
that $\beta = \alpha + 2q$. If we were to now apply the midpoint rule to this subinterval, then we would write
\[
\int_{\alpha}^{\beta} f(x) \, dx \approx 2q \cdot f(\alpha + q) = qf(\alpha + q) + qf(\beta - q)
\]
since the interval has width $2q$ and the midpoint is $\alpha + q = \beta - q$.

The sneaky trick we will employ is to write
\[
\int_{\alpha}^{\beta} f(x) \, dx = \int_{\alpha}^{\alpha + q} f(x) \, dx + \int_{\alpha + q}^{\beta} f(x) \, dx
\]
and then examine each of the integrals on the right-hand side (using integration by parts) and show that they are each of the form
\[
\int_{\alpha}^{\alpha + q} f(x) \, dx \approx qf(\alpha + q) + \text{small error term}
\]
\[
\int_{\alpha + q}^{\beta} f(x) \, dx \approx qf(\beta - q) + \text{small error term}
\]

Let us apply integration by parts to $\int_{\alpha}^{\alpha + q} f(x) \, dx$ — with $u = f(x), dv = dx$ so $du = f'(x) \, dx$ and we will make the slightly non-standard choice of $v = x - \alpha$:
\[
\int_{\alpha}^{\alpha + q} f(x) \, dx = \left[(x - \alpha)f(x)\right]_{\alpha}^{\alpha + q} - \int_{\alpha}^{\alpha + q} (x - \alpha)f'(x) \, dx
\]
\[
= qf(\alpha + q) - \int_{\alpha}^{\alpha + q} (x - \alpha)f'(x) \, dx
\]
Notice that the first term on the right-hand side is the term we need, and that our non-standard choice of $v$ allowed us to avoid introducing an $f(\alpha)$ term.

Now integrate by parts again using $u = f'(x), dv = (x - \alpha) \, dx$, so $du = f''(x), v = \frac{(x-\alpha)^2}{2}$:
\[
\int_{\alpha}^{\alpha + q} f(x) \, dx = qf(\alpha + q) - \int_{\alpha}^{\alpha + q} (x - \alpha)f'(x) \, dx
\]
\[
= qf(\alpha + q) - \left[\frac{(x-\alpha)^2}{2}f'(x)\right]_{\alpha}^{\alpha + q} + \int_{\alpha}^{\alpha + q} (x - \alpha)^2f''(x) \, dx
\]
\[
= qf(\alpha + q) - \frac{q^2}{2}f'(\alpha + q) + \int_{\alpha}^{\alpha + q} (x - \alpha)^2f''(x) \, dx
\]

To obtain a similar expression for the other integral, we repeat the above steps and obtain:
\[
\int_{\alpha + q}^{\beta} f(x) \, dx = qf(\beta - q) + \frac{q^2}{2}f'(\beta - q) + \int_{\beta - q}^{\beta} (x - \beta)^2f''(x) \, dx
\]
Now add together these two expressions
\[
\int_{\alpha}^{\alpha+q} f(x) \, dx + \int_{\beta-q}^{\beta} f(x) \, dx = qf(\alpha + q) + qf(\beta - q) + \frac{q^2}{2} (f'(\beta - q) - f'(\alpha + q)) + \int_{\alpha}^{\alpha+q} \frac{(x - \alpha)^2}{2} f''(x) \, dx + \int_{\beta-q}^{\beta} \frac{(x - \beta)^2}{2} f''(x) \, dx
\]
Then since \( \alpha + q = \beta - q \) we can combine the integrals on the left-hand side and eliminate some terms from the right-hand side:
\[
\int_{\alpha}^{\beta} f(x) \, dx = 2qf(\alpha + q) + \int_{\alpha}^{\alpha+q} \frac{(x - \alpha)^2}{2} f''(x) \, dx + \int_{\beta-q}^{\beta} \frac{(x - \beta)^2}{2} f''(x) \, dx
\]
Rearrange this expression a little and take absolute values
\[
\left| \int_{\alpha}^{\beta} f(x) \, dx - 2qf(\alpha + q) \right| \leq \left| \int_{\alpha}^{\alpha+q} \frac{(x - \alpha)^2}{2} f''(x) \, dx \right| + \left| \int_{\beta-q}^{\beta} \frac{(x - \beta)^2}{2} f''(x) \, dx \right|
\]
where we have also made use of the triangle inequality. By assumption \(|f''(x)| \leq M\) on the interval \(\alpha \leq x \leq \beta\), so
\[
\left| \int_{\alpha}^{\beta} f(x) \, dx - 2qf(\alpha + q) \right| \leq M \int_{\alpha}^{\alpha+q} \frac{(x - \alpha)^2}{2} \, dx + M \int_{\beta-q}^{\beta} \frac{(x - \beta)^2}{2} \, dx
\]
\[
= \frac{Mq^3}{3} = \frac{M(\beta - \alpha)^3}{24}
\]
where we have used \(q = \frac{\beta - \alpha}{2}\) in the last step.

Thus on any interval \(x_i \leq x \leq x_{i+1} = x_i + \Delta x\)
\[
\left| \int_{x_i}^{x_{i+1}} f(x) \, dx - \Delta x f \left( \frac{x_i + x_{i+1}}{2} \right) \right| \leq \frac{M}{24} (\Delta x)^3
\]
Putting everything together we see that the error using the midpoint rule is bounded by
\[
\left| \int_{\alpha}^{\beta} f(x) \, dx - [f(\bar{x}_1) + f(\bar{x}_2) + \cdots + f(\bar{x}_n)] \Delta x \right|
\]
\[
\leq \left| \int_{x_0}^{x_1} f(x) \, dx - \Delta x f(\bar{x}_1) \right| + \cdots + \left| \int_{x_{n-1}}^{x_n} f(x) \, dx - \Delta x f(\bar{x}_n) \right|
\]
\[
\leq n \times \frac{M}{24} (\Delta x)^3 = n \times \frac{M}{24} \left( \frac{b - a}{n} \right)^3 = \frac{M(b - a)^3}{24n^2}
\]
as required.

A very similar analysis shows that, as was stated in Theorem 1.11.12 above,

83 The triangle inequality says that for any real numbers \(x, y\)
\[
|x + y| \leq |x| + |y|.
\]
the total error introduced by the trapezoidal rule is bounded by \( \frac{M}{12} \frac{(b-a)^3}{n^2} \),

- the total error introduced by Simpson’s rule is bounded by \( \frac{M}{180} \frac{(b-a)^5}{n^4} \).

1.12 Improper Integrals

1.12.1 Definitions

To this point we have only considered nicely behaved integrals \( \int_a^b f(x) \, dx \). Though the algebra involved in some of our examples was quite difficult, all the integrals had

- finite limits of integration \( a \) and \( b \), and

- a bounded integrand \( f(x) \) (and in fact continuous except possibly for finitely many jump discontinuities).

Not all integrals we need to study are quite so nice.

**Definition 1.12.1.**

An integral having either an infinite limit of integration or an unbounded integrand is called an improper integral.

Two examples are

\[
\int_0^\infty \frac{dx}{1 + x^2} \quad \text{and} \quad \int_0^1 \frac{dx}{x}
\]

The first has an infinite domain of integration and the integrand of the second tends to \( \infty \) as \( x \) approaches the left end of the domain of integration. We’ll start with an example that illustrates the traps that you can fall into if you treat such integrals sloppily. Then we’ll see how to treat them carefully.

**Example 1.12.2 \( \int_{-1}^1 \frac{1}{x^2} \, dx \)**

Consider the integral

\[
\int_{-1}^1 \frac{1}{x^2} \, dx
\]

If we “do” this integral completely naively then we get

\[
\int_{-1}^1 \frac{1}{x^2} \, dx = \left. \frac{x^{-1}}{-1} \right|_{-1}^{1}
\]

\[
= \frac{1}{-1} - \frac{-1}{-1}
\]

\[
= -2
\]
which is *wrong*. In fact, the answer is ridiculous. The integrand \( \frac{1}{x^2} > 0 \), so the integral has to be positive.

The flaw in the argument is that the fundamental theorem of calculus, which says that

\[
\text{if } F'(x) = f(x) \text{ then } \int_a^b f(x) \, dx = F(b) - F(a)
\]

is applicable only when \( F'(x) \) exists and equals \( f(x) \) for all \( a \leq x \leq b \). In this case \( F'(x) = \frac{1}{x^2} \) does not exist for \( x = 0 \). The given integral is improper. We’ll see later that the correct answer is \( +\infty \).

Example 1.12.2

Let us put this example to one side for a moment and turn to the integral \( \int_a^\infty \frac{dx}{1+x^2} \). In this case, the integrand is bounded but the domain of integration extends to \( +\infty \). We can evaluate this integral by sneaking up on it. We compute it on a bounded domain of integration, like \( \int_a^R \frac{dx}{1+x^2} \), and then take the limit \( R \to \infty \). Let us put this into practice:

\[
\begin{align*}
\int_a^R \frac{dx}{1+x^2} &= \arctan x \bigg|_a^R \\
&= \arctan R - \arctan a
\end{align*}
\]

Example 1.12.3 \( \left( \int_a^\infty \frac{dx}{1+x^2} \right) \)

Solution.

- Since the domain extends to \( +\infty \) we first integrate on a finite domain
  \[
  \int_a^R \frac{dx}{1+x^2} = \arctan x \bigg|_a^R = \arctan R - \arctan a
  \]

- We then take the limit as \( R \to +\infty \):
  \[
  \int_a^\infty \frac{dx}{1+x^2} = \lim_{R \to \infty} \int_a^R \frac{dx}{1+x^2} = \lim_{R \to \infty} \left[ \arctan R - \arctan a \right] = \frac{\pi}{2} - \arctan a.
  \]

84 Very wrong. But it is not an example of “not even wrong” — which is a phrase attributed to the physicist Wolfgang Pauli who was known for his harsh critiques of sloppy arguments. The phrase is typically used to describe arguments that are so incoherent that not only can one not prove they are true, but they lack enough coherence to be able to show they are false. The interested reader should do a little searchengineering and look at the concept of falsifyability.
To be more precise, we actually formally define an integral with an infinite domain as the limit of the integral with a finite domain as we take one or more of the limits of integration to infinity.

**Definition 1.12.4 (Improper integral with infinite domain of integration).**

(a) If the integral \( \int_a^R f(x) \, dx \) exists for all \( R > a \), then
\[
\int_a^\infty f(x) \, dx = \lim_{R \to \infty} \int_a^R f(x) \, dx
\]
when the limit exists (and is finite).

(b) If the integral \( \int_r^b f(x) \, dx \) exists for all \( r < b \), then
\[
\int_{-\infty}^b f(x) \, dx = \lim_{r \to -\infty} \int_r^b f(x) \, dx
\]
when the limit exists (and is finite).

(c) If the integral \( \int_r^R f(x) \, dx \) exists for all \( r < R \), then
\[
\int_{-\infty}^\infty f(x) \, dx = \lim_{r \to -\infty} \int_r^c f(x) \, dx + \lim_{R \to \infty} \int_c^R f(x) \, dx
\]
when both limits exist (and are finite). Any \( c \) can be used.

When the limit(s) exist, the integral is said to be convergent. Otherwise it is said to be divergent.

We must also be able to treat an integral like \( \frac{1}{x} \) that has a finite domain of integration but whose integrand is unbounded near one limit of integration. Our approach is similar — we sneak up on the problem. We compute the integral on a smaller domain, such as \( \int_t^1 \frac{dx}{x} \), with \( t > 0 \), and then take the limit \( t \to 0^+ \).

**Example 1.12.5 \( \int_0^1 \frac{1}{x} \, dx \)**

**Solution.**

- Since the integrand is unbounded near \( x = 0 \), we integrate on the smaller domain

\[ 85 \text{ This will, in turn, allow us to deal with integrals whose integrand is unbounded somewhere inside the domain of integration.} \]
$t \leq x \leq 1$ with $t > 0$:

$$\int_{t}^{1} \frac{1}{x} \, dx = \log |x| \bigg|_{t}^{1} = -\log |t|$$

- We then take the limit as $t \to 0^+$ to obtain

$$\int_{0}^{1} \frac{1}{x} \, dx = \lim_{t \to 0^+} \int_{t}^{1} \frac{1}{x} \, dx = \lim_{t \to 0^+} -\log |t| = +\infty$$

Thus this integral diverges to $+\infty$.

Indeed, we define integrals with unbounded integrands via this process:
(a) If the integral \( \int_a^b f(x) \, dx \) exists for all \( a < t < b \), then
\[
\int_a^b f(x) \, dx = \lim_{t \to a^+} \int_t^b f(x) \, dx
\]
when the limit exists (and is finite).

(b) If the integral \( \int_a^T f(x) \, dx \) exists for all \( a < T < b \), then
\[
\int_a^b f(x) \, dx = \lim_{T \to b^-} \int_a^T f(x) \, dx
\]
when the limit exists (and is finite).

(c) Let \( a < c < b \). If the integrals \( \int_a^T f(x) \, dx \) and \( \int_t^b f(x) \, dx \) exist for all \( a < T < c \) and \( c < t < b \), then
\[
\int_a^b f(x) \, dx = \lim_{T \to c^-} \int_a^T f(x) \, dx + \lim_{t \to c^+} \int_t^b f(x) \, dx
\]
when both limit exist (and are finite).

When the limit(s) exist, the integral is said to be convergent. Otherwise it is said to be divergent.

Notice that (c) is used when the integrand is unbounded at some point in the middle of the domain of integration, such as was the case in our original example
\[
\int_{-1}^1 \frac{1}{x^2} \, dx
\]
A quick computation shows that this integral diverges to \(+\infty\)
\[
\int_{-1}^1 \frac{1}{x^2} \, dx = \lim_{a \to 0^-} \int_{-1}^a \frac{1}{x^2} \, dx + \lim_{b \to 0^+} \int_b^1 \frac{1}{x^2} \, dx
\]
\[
= \lim_{a \to 0^-} \left[ 1 - \frac{1}{a} \right] + \lim_{b \to 0^+} \left[ \frac{1}{b} - 1 \right]
\]
\[= +\infty \]

More generally, if an integral has more than one “source of impropriety” (for example an infinite domain of integration and an integrand with an unbounded integrand or multiple infinite discontinuities) then you split it up into a sum of integrals with a single “source of impropriety” in each. For the integral, as a whole, to converge every term in that sum has to converge.

For example
Example 1.12.7 \( \left( \int_{-\infty}^{\infty} \frac{dx}{(x-2)x^2} \right) \)

Consider the integral

\[ \int_{-\infty}^{\infty} \frac{dx}{(x-2)x^2} \]

- The domain of integration that extends to both \( +\infty \) and \( -\infty \).
- The integrand is singular (i.e. becomes infinite) at \( x = 2 \) and at \( x = 0 \).
- So we would write the integral as

\[ \int_{-\infty}^{\infty} \frac{dx}{(x-2)x^2} = \int_{-\infty}^{a} \frac{dx}{(x-2)x^2} + \int_{a}^{0} \frac{dx}{(x-2)x^2} + \int_{0}^{b} \frac{dx}{(x-2)x^2} + \int_{b}^{2} \frac{dx}{(x-2)x^2} + \int_{2}^{c} \frac{dx}{(x-2)x^2} + \int_{c}^{\infty} \frac{dx}{(x-2)x^2} \]

where

- \( a \) is any number strictly less than 0,
- \( b \) is any number strictly between 0 and 2, and
- \( c \) is any number strictly bigger than 2.

So, for example, take \( a = -1, b = 1, c = 3 \).

- When we examine the right-hand side we see that
  - the first integral has domain of integration extending to \( -\infty \)
  - the second integral has an integrand that becomes unbounded as \( x \to 0^- \),
  - the third integral has an integrand that becomes unbounded as \( x \to 0^+ \),
  - the fourth integral has an integrand that becomes unbounded as \( x \to 2^- \),
  - the fifth integral has an integrand that becomes unbounded as \( x \to 2^+ \), and
  - the last integral has domain of integration extending to \( +\infty \).

- Each of these integrals can then be expressed as a limit of an integral on a small domain.

1.12.2 Examples

With the more formal definitions out of the way, we are now ready for some (important) examples.

Example 1.12.8 \( \left( \int_{1}^{\infty} \frac{dx}{x^p} \text{ with } p > 0 \right) \)

Solution.
• Fix any \( p > 0 \).

• The domain of the integral \( \int_1^{\infty} \frac{dx}{x^p} \) extends to \( +\infty \) and the integrand \( \frac{1}{x^p} \) is continuous and bounded on the whole domain.

• So we write this integral as the limit

\[
\int_1^{\infty} \frac{dx}{x^p} = \lim_{R \to \infty} \int_1^{R} \frac{dx}{x^p}
\]

• The antiderivative of \( \frac{1}{x^p} \) changes when \( p = 1 \), so we will split the problem into three cases, \( p > 1 \), \( p = 1 \) and \( p < 1 \).

• When \( p > 1 \),

\[
\int_1^{R} \frac{dx}{x^p} = \frac{1}{1-p} \left( x^{1-p} \right) \bigg|_1^R = \frac{R^{1-p} - 1}{1-p}
\]

Taking the limit as \( R \to \infty \) gives

\[
\int_1^{\infty} \frac{dx}{x^p} = \lim_{R \to \infty} \int_1^{R} \frac{dx}{x^p} = \lim_{R \to \infty} \frac{R^{1-p} - 1}{1-p}
\]

\[
= -\frac{1}{1-p} = \frac{1}{p-1}
\]

since \( 1 - p < 0 \).

• Similarly when \( p < 1 \) we have

\[
\int_1^{\infty} \frac{dx}{x^p} = \lim_{R \to \infty} \int_1^{R} \frac{dx}{x^p} = \lim_{R \to \infty} \frac{R^{1-p} - 1}{1-p}
\]

because \( 1 - p > 0 \) and the term \( R^{1-p} \) diverges to \( +\infty \).

• Finally when \( p = 1 \)

\[
\int_1^{R} \frac{dx}{x} = \log |R| - \log 0 = \log R
\]

Then taking the limit as \( R \to \infty \) gives us

\[
\int_1^{\infty} \frac{dx}{x^p} = \lim_{R \to \infty} \log |R| = +\infty.
\]
• So summarising, we have

\[
\int_1^\infty \frac{dx}{x^p} = \begin{cases} 
\text{divergent} & \text{if } p \leq 1 \\
\frac{1}{p-1} & \text{if } p > 1 
\end{cases}
\]

Example 1.12.8

Example 1.12.9 \( \int_0^1 \frac{dx}{x^p} \) with \( p > 0 \)

Solution.

• Again fix any \( p > 0 \).

• The domain of integration of the integral \( \int_0^1 \frac{dx}{x^p} \) is finite, but the integrand \( \frac{1}{x^p} \) becomes unbounded as \( x \) approaches the left end, 0, of the domain of integration.

• So we write this integral as

\[
\int_0^1 \frac{dx}{x^p} = \lim_{t \to 0^+} \int_t^1 \frac{dx}{x^p}
\]

• Again, the antiderivative changes at \( p = 1 \), so we split the problem into three cases.

• When \( p > 1 \) we have

\[
\int_t^1 \frac{dx}{x^p} = \frac{1}{1-p} x^{1-p} \bigg|_t^1 \\
= \frac{1 - t^{1-p}}{1-p}
\]

Since \( 1 - p < 0 \) when we take the limit as \( t \to 0 \) the term \( t^{1-p} \) diverges to \( +\infty \) and we obtain

\[
\int_0^1 \frac{dx}{x^p} = \lim_{t \to 0^+} \frac{1 - t^{1-p}}{1-p} = +\infty
\]

• When \( p = 1 \) we similarly obtain

\[
\int_0^1 \frac{dx}{x} = \lim_{t \to 0^+} \int_t^0 \frac{dx}{x} \\
= \lim_{t \to 0^+} (- \log |t|) \\
= +\infty
\]
Finally, when $p < 1$ we have

$$
\int_0^1 \frac{dx}{x^p} = \lim_{t \to 0^+} \int_t^1 \frac{dx}{x^p} = \lim_{t \to 0^+} \frac{1 - t^{1-p}}{1 - p} = \frac{1}{1 - p}
$$

since $1 - p > 0$.

In summary

$$
\int_0^1 \frac{dx}{x^p} = \begin{cases} 
\frac{1}{1-p} & \text{if } p < 1 \\
\text{divergent} & \text{if } p \geq 1
\end{cases}
$$

Example 1.12.9

Example 1.12.10 \( \int_0^\infty \frac{dx}{x^p} \) with $p > 0$

Solution.

Yet again fix $p > 0$.

This time the domain of integration of the integral $\int_0^\infty \frac{dx}{x^p}$ extends to $+\infty$, and in addition the integrand $\frac{1}{x^p}$ becomes unbounded as $x$ approaches the left end, 0, of the domain of integration.

So we split the domain in two — given our last two examples, the obvious place to cut is at $x = 1$:

$$
\int_0^\infty \frac{dx}{x^p} = \int_0^1 \frac{dx}{x^p} + \int_1^\infty \frac{dx}{x^p}
$$

We saw, in Example 1.12.9, that the first integral diverged whenever $p \geq 1$, and we also saw, in Example 1.12.8, that the second integral diverged whenever $p \leq 1$.

So the integral $\int_0^\infty \frac{dx}{x^p}$ diverges for all values of $p$.

Example 1.12.10

Example 1.12.11 \( \int_{-1}^1 \frac{dx}{x} \)

This is a pretty subtle example. Look at the sketch below: This suggests that the signed
area to the left of the $y$–axis should exactly cancel the area to the right of the $y$–axis making the value of the integral $\int_{-1}^{1} \frac{dx}{x}$ exactly zero.

But both of the integrals

\[
\int_{0}^{1} \frac{dx}{x} = \lim_{t \to 0^+} \int_{0}^{1} \frac{dx}{x} = \lim_{t \to 0^+} \left[ \log x \right]_{t}^{1} = \lim_{t \to 0^+} \log \frac{1}{t} = +\infty
\]

\[
\int_{-1}^{0} \frac{dx}{x} = \lim_{T \to 0^-} \int_{-1}^{T} \frac{dx}{x} = \lim_{T \to 0^-} \left[ \log |x| \right]_{-1}^{T} = \lim_{T \to 0^-} \log |T| = -\infty
\]

diverge so $\int_{-1}^{1} \frac{dx}{x}$ diverges. Don’t make the mistake of thinking that $\infty - \infty = 0$. It is undefined. And it is undefined for good reason.

For example, we have just seen that the area to the right of the $y$–axis is

\[
\lim_{t \to 0^+} \int_{t}^{1} \frac{dx}{x} = +\infty
\]

and that the area to the left of the $y$–axis is (substitute $-7t$ for $T$ above)

\[
\lim_{t \to 0^+} \int_{-1}^{-7t} \frac{dx}{x} = -\infty
\]

If $\infty - \infty = 0$, the following limit should be 0.

\[
\lim_{t \to 0^+} \left[ \int_{t}^{1} \frac{dx}{x} + \int_{-1}^{-7t} \frac{dx}{x} \right] = \lim_{t \to 0^+} \left[ \log \frac{1}{t} + \log |-7t| \right]
\]

\[
= \lim_{t \to 0^+} \left[ \log \frac{1}{t} + \log(7t) \right]
\]

\[
= \lim_{t \to 0^+} \left[ -\log t + \log 7 + \log t \right] = \lim_{t \to 0^+} \log 7
\]

This appears to give $\infty - \infty = \log 7$. Of course the number 7 was picked at random. You can make $\infty - \infty$ be any number at all, by making a suitable replacement for 7.
Example 1.12.11

Example 1.12.12 (Example 1.12.2 revisited)

The careful computation of the integral of Example 1.12.2 is

\[
\int_{-1}^{1} \frac{1}{x^2} \, dx = \lim_{T \to 0^-} \int_{-1}^{T} \frac{1}{x^2} \, dx + \lim_{t \to 0^+} \int_{t}^{1} \frac{1}{x^2} \, dx \\
= \lim_{T \to 0^-} \left[ - \frac{1}{x} \right]_{-1}^{T} + \lim_{t \to 0^+} \left[ - \frac{1}{x} \right]_{t}^{1} \\
= \infty + \infty
\]

Hence the integral diverges to \(+\infty\).

Example 1.12.12

Example 1.12.13 \(\int_{-\infty}^{\infty} \frac{dx}{1+x^2}\)

Since

\[
\lim_{R \to \infty} \int_{0}^{R} \frac{dx}{1 + x^2} = \lim_{R \to \infty} \left[ \arctan x \right]_{0}^{R} = \lim_{R \to \infty} \arctan R = \frac{\pi}{2} \\
\lim_{r \to -\infty} \int_{r}^{0} \frac{dx}{1 + x^2} = \lim_{r \to -\infty} \left[ \arctan x \right]_{r}^{0} = \lim_{r \to -\infty} -\arctan r = \frac{\pi}{2}
\]

The integral \(\int_{-\infty}^{\infty} \frac{dx}{1+x^2}\) converges and takes the value \(\pi\).

Example 1.12.13

Example 1.12.14

For what values of \(p\) does \(\int_{e}^{\infty} \frac{dx}{x(\log x)^p}\) converge?

Solution.

- For \(x \geq e\), the denominator \(x(\log x)^p\) is never zero. So the integrand is bounded on the entire domain of integration and this integral is improper only because the domain of integration extends to \(+\infty\) and we proceed as usual.
We have
\[ \int_e^\infty \frac{dx}{x (\log x)^p} = \lim_{R \to \infty} \int_e^R \frac{dx}{x (\log x)^p} \]
use substitution
\[ = \lim_{R \to \infty} \int_1^{\log R} \frac{du}{u^p} \]
with \( u = \log x, \frac{du}{x} = \frac{dx}{x} \)

\[ = \lim_{R \to \infty} \begin{cases} \frac{1}{1-p} \left[ (\log R)^{1-p} - 1 \right] & \text{if } p \neq 1 \\ \log(\log R) & \text{if } p = 1 \end{cases} \]

\[ = \begin{cases} \text{divergent} & \text{if } p \leq 1 \\ \frac{1}{p-1} & \text{if } p > 1 \end{cases} \]

In this last step we have used similar logic that that used in Example 1.12.8, but with \( R \) replaced by \( \log R \).

---

Example 1.12.15 (the gamma function)

The gamma function \( \Gamma(x) \) is defined by the improper integral
\[ \Gamma(t) = \int_0^\infty x^{t-1} e^{-x} \, dx \]

We shall now compute \( \Gamma(n) \) for all natural numbers \( n \).

- To get started, we’ll compute
\[ \Gamma(1) = \int_0^\infty e^{-x} \, dx = \lim_{R \to \infty} \int_0^R e^{-x} \, dx = \lim_{R \to \infty} \left[ -e^{-x} \right]_0^R = 1 \]

- Then compute
\[ \Gamma(2) = \int_0^\infty xe^{-x} \, dx \]
use integration by parts with \( u = x, \frac{dv}{dx} = e^{-x} \, dx, \)
\[ v = -e^{-x}, \frac{du}{dx} = dx \]

\[ = \lim_{R \to \infty} \left[ -xe^{-x} \right]_0^R + \int_0^R e^{-x} \, dx \]
\[ = \lim_{R \to \infty} \left[ -xe^{-x} - e^{-x} \right]_0^R \]
\[ = 1 \]

For the last equality, we used that \( \lim_{x \to \infty} xe^{-x} = 0 \).
Now we move on to general \( n \), using the same type of computation as we just used to evaluate \( \Gamma(2) \). For any natural number \( n \),

\[
\Gamma(n + 1) = \int_0^\infty x^n e^{-x} \, dx
\]

\[
= \lim_{R \to \infty} \int_0^R x^n e^{-x} \, dx
\]

again integrate by parts with

\[
u = x^n, \quad dv = e^{-x} \, dx,
\]

\[
v = -e^{-x}, \quad du = nx^{n-1} \, dx
\]

\[
= \lim_{R \to \infty} \left[ -x^n e^{-x}\bigg|_0^R + \int_0^R nx^{n-1} e^{-x} \, dx \right]
\]

\[
= \lim_{R \to \infty} n \int_0^R x^{n-1} e^{-x} \, dx
\]

\[
= n \Gamma(n)
\]

To get to the third row, we used that \( \lim_{x \to \infty} x^n e^{-x} = 0 \).

Now that we know \( \Gamma(2) = 1 \) and \( \Gamma(n + 1) = n \Gamma(n) \), for all \( n \in \mathbb{N} \), we can compute all of the \( \Gamma(n) \)'s.

\[
\Gamma(2) = 1
\]
\[
\Gamma(3) = \Gamma(2 + 1) = 2 \Gamma(2) = 2 \cdot 1
\]
\[
\Gamma(4) = \Gamma(3 + 1) = 3 \Gamma(3) = 3 \cdot 2 \cdot 1
\]
\[
\Gamma(5) = \Gamma(4 + 1) = 4 \Gamma(4) = 4 \cdot 3 \cdot 2 \cdot 1
\]

\[
\vdots
\]

\[
\Gamma(n) = (n - 1) \cdot (n - 2) \cdot \cdots \cdot 4 \cdot 3 \cdot 2 \cdot 1 = (n - 1)!
\]

That is, the factorial is just\(^{86}\) the Gamma function shifted by one.

\[\text{Example 1.12.15}\]

### 1.12.3 Convergence Tests for Improper Integrals

It is very common to encounter integrals that are too complicated to evaluate explicitly. Numerical approximation schemes, evaluated by computer, are often used instead (see Section 1.11). You want to be sure that at least the integral converges before feeding it into

\[\text{Example 1.12.15}\]

\[\begin{equation}
\Gamma(1 - z) \Gamma(z) = \frac{\pi}{\sin \pi z}.
\end{equation}\]

\[\text{Example 1.12.15}\]

---

\(^{86}\) The Gamma function is far more important than just a generalisation of the factorial. It appears all over mathematics, physics, statistics and beyond. It has all sorts of interesting properties and its definition can be extended from natural numbers \( n \) to all numbers excluding 0, \(-1, -2, -3, \ldots\). For example, one can show that

\[
\Gamma(1 - z) \Gamma(z) = \frac{\pi}{\sin \pi z}.
\]
a computer. Fortunately it is usually possible to determine whether or not an improper integral converges even when you cannot evaluate it explicitly.

**Remark 1.12.16.** For pedagogical purposes, we are going to concentrate on the problem of determining whether or not an integral $\int_{a}^{\infty} f(x) \, dx$ converges, when $f(x)$ has no singularities for $x \geq a$. Recall that the first step in analyzing any improper integral is to write it as a sum of integrals each of has only a single “source of impropriety” — either a domain of integration that extends to $+\infty$, or a domain of integration that extends to $-\infty$, or an integrand which is singular at one end of the domain of integration. So we are now going to consider only the first of these three possibilities. But the techniques that we are about to see have obvious analogues for the other two possibilities.

Now let’s start. Imagine that we have an improper integral $\int_{a}^{\infty} f(x) \, dx$, that $f(x)$ has no singularities for $x \geq a$ and that $f(x)$ is complicated enough that we cannot evaluate the integral explicitly. The idea is find another improper integral $\int_{a}^{\infty} g(x) \, dx$

- with $g(x)$ simple enough that we can evaluate the integral $\int_{a}^{\infty} g(x) \, dx$ explicitly, or at least determine easily whether or not $\int_{a}^{\infty} g(x) \, dx$ converges, and

- with $g(x)$ behaving enough like $f(x)$ for large $x$ that the integral $\int_{a}^{\infty} f(x) \, dx$ converges if and only if $\int_{a}^{\infty} g(x) \, dx$ converges.

So far, this is a pretty vague strategy. Here is a theorem which starts to make it more precise.

**Theorem 1.12.17 (Comparison).**

Let $a$ be a real number. Let $f$ and $g$ be functions that are defined and continuous for all $x \geq a$ and assume that $g(x) \geq 0$ for all $x \geq 0$.

(a) If $|f(x)| \leq g(x)$ for all $x \geq a$ and if $\int_{a}^{\infty} g(x) \, dx$ converges then $\int_{a}^{\infty} f(x) \, dx$ also converges.

(b) If $f(x) \geq g(x)$ for all $x \geq a$ and if $\int_{a}^{\infty} g(x) \, dx$ diverges then $\int_{a}^{\infty} f(x) \, dx$ diverges.

We will not prove this theorem, but, hopefully, the following supporting arguments should at least appear reasonable to you. Consider the figure below:

---

87 Applying numerical integration methods to a divergent integral may result in perfectly reasonably looking but very wrong answers.

88 You could, for example, think of something like our running example $\int_{a}^{\infty} e^{-t^2} \, dt$. 

200
• If \( \int_a^\infty g(x) \, dx \) converges, then the area of
  \[
  \{(x,y) \mid x \geq a, \, 0 \leq y \leq g(x)\}
  \]
is finite.

When \( |f(x)| \leq g(x) \), the region
  \[
  \{(x,y) \mid x \geq a, \, 0 \leq y \leq |f(x)|\}
  \]
is contained inside \( \{(x,y) \mid x \geq a, \, 0 \leq y \leq g(x)\} \) and so must also have finite area. Consequently the areas of both the regions
  \[
  \{(x,y) \mid x \geq a, \, 0 \leq y \leq f(x)\}
  \]
and \( \{(x,y) \mid x \geq a, \, f(x) \leq y \leq 0\} \) are finite too. \(^{89}\)

• If \( \int_a^\infty g(x) \, dx \) diverges, then the area of
  \[
  \{(x,y) \mid x \geq a, \, 0 \leq y \leq g(x)\}
  \]
is infinite.

When \( f(x) \geq g(x) \), the region
  \[
  \{(x,y) \mid x \geq a, \, 0 \leq y \leq f(x)\}
  \]
contains the region \( \{(x,y) \mid x \geq a, \, 0 \leq y \leq g(x)\} \) and so also has infinite area.

### Example 1.12.18 \( \int_1^\infty e^{-x^2} \, dx \)

We cannot evaluate the integral \( \int_1^\infty e^{-x^2} \, dx \) explicitly,\(^{90}\) however we would still like to understand if it is finite or not — does it converge or diverge?

**Solution.** We will use Theorem 1.12.17 to answer the question.

• So we want to find another integral that we can compute and that we can compare to \( \int_1^\infty e^{-x^2} \, dx \). To do so we pick an integrand that looks like \( e^{-x^2} \), but whose indefinite integral we know — such as \( e^{-x} \).

• When \( x \geq 1 \), we have \( x^2 \geq x \) and hence \( e^{-x^2} \leq e^{-x} \). Thus we can use Theorem 1.12.17 to compare
  \[
  \int_1^\infty e^{-x^2} \, dx \text{ with } \int_1^\infty e^{-x} \, dx
  \]
• The integral
  \[
  \int_1^\infty e^{-x} \, dx = \lim_{R \to \infty} \int_1^R e^{-x} \, dx
  \]
  \[
  = \lim_{R \to \infty} \left[ -e^{-x} \right]_1^R
  \]
  \[
  = \lim_{R \to \infty} \left[ e^{-1} - e^{-R} \right] = e^{-1}
  \]
converges.

\(^{89}\) We have separated the regions in which \( f(x) \) is positive and negative, because the integral \( \int_a^\infty f(x) \, dx \) represents the signed area of the union of \( \{(x,y) \mid x \geq a, \, 0 \leq y \leq f(x)\} \) and \( \{(x,y) \mid x \geq a, \, f(x) \leq y \leq 0\} \).

\(^{90}\) It has been the subject of many remarks and footnotes.
So, by Theorem 1.12.17, with \( a = 1 \), \( f(x) = e^{-x^2} \) and \( g(x) = e^{-x} \), the integral \( \int_{1}^{\infty} e^{-x^2} \, dx \) converges too (it is approximately equal to 0.1394).

Example 1.12.18

**Example 1.12.19**  \( \int_{1/2}^{\infty} e^{-x^2} \, dx \)

**Solution.**

- The integral \( \int_{1/2}^{\infty} e^{-x^2} \, dx \) is quite similar to the integral \( \int_{1}^{\infty} e^{-x^2} \, dx \) of Example 1.12.18. But we cannot just repeat the argument of Example 1.12.18 because it is not true that \( e^{-x^2} \leq e^{-x} \) when \( 0 < x < 1 \).

- In fact, for \( 0 < x < 1 \), \( x^2 < x \) so that \( e^{-x^2} > e^{-x} \).

- However the difference between the current example and Example 1.12.18 is

\[
\int_{1/2}^{\infty} e^{-x^2} \, dx - \int_{1}^{\infty} e^{-x^2} \, dx = \int_{1/2}^{1} e^{-x^2} \, dx
\]

which is clearly a well defined finite number (its actually about 0.286). It is important to note that we are being a little sloppy by taking the difference of two integrals like this — we are assuming that both integrals converge. More on this below.

- So we would expect that \( \int_{1/2}^{\infty} e^{-x^2} \, dx \) should be the sum of the proper integral integral \( \int_{1/2}^{1} e^{-x^2} \, dx \) and the convergent integral \( \int_{1}^{\infty} e^{-x^2} \, dx \) and so should be a convergent integral. This is indeed the case. The Theorem below provides the justification.

**Theorem 1.12.20.**

Let \( a \) and \( c \) be real numbers with \( a < c \) and let the function \( f(x) \) be continuous for all \( x \geq a \). Then the improper integral \( \int_{a}^{\infty} f(x) \, dx \) converges if and only if the improper integral \( \int_{c}^{\infty} f(x) \, dx \) converges.

**Proof.** By definition the improper integral \( \int_{a}^{\infty} f(x) \, dx \) converges if and only if the limit

\[
\lim_{R \to \infty} \int_{a}^{R} f(x) \, dx = \lim_{R \to \infty} \left[ \int_{a}^{c} f(x) \, dx + \int_{c}^{R} f(x) \, dx \right]
= \int_{a}^{c} f(x) \, dx + \lim_{R \to \infty} \int_{c}^{R} f(x) \, dx
\]
exists and is finite. (Remember that, in computing the limit, \( \int_a^c f(x) \, dx \) is a finite constant independent of \( R \) and so can be pulled out of the limit.) But that is the case if and only if the limit \( \lim_{R \to \infty} \int_c^R f(x) \, dx \) exists and is finite, which in turn is the case if and only if the integral \( \int_c^\infty f(x) \, dx \) converges.

**Example 1.12.21**

Does the integral \( \int_1^\infty \frac{\sqrt{x}}{x^2+x} \, dx \) converge or diverge?

**Solution.**

- Our first task is to identify the potential sources of impropriety for this integral.
- The domain of integration extends to \( +\infty \), but we must also check to see if the integrand contains any singularities. On the domain of integration \( x \geq 1 \) so the denominator is never zero and the integrand is continuous. So the only problem is at \( +\infty \).
- Our second task is to develop some intuition\(^91\). As the only problem is that the domain of integration extends to infinity, whether or not the integral converges will be determined by the behavior of the integrand for very large \( x \).
- When \( x \) is very large, \( x^2 \) is much much larger than \( x \) (which we can write as \( x^2 \gg x \)) so that the denominator \( x^2 + x \approx x^2 \) and the integrand
  \[
  \frac{\sqrt{x}}{x^2+x} \approx \frac{\sqrt{x}}{x^2} = \frac{1}{x^{3/2}}
  \]
- By Example 1.12.8, with \( p = 3/2 \), the integral \( \int_1^\infty \frac{dx}{x^{3/2}} \) converges. So we would expect that \( \int_1^\infty \frac{\sqrt{x}}{x^2+x} \, dx \) converges too.
- Our final task is to verify that our intuition is correct. To do so, we want to apply part (a) of Theorem 1.12.17 with \( f(x) = \frac{\sqrt{x}}{x^2+x} \) and \( g(x) \) being \( \frac{1}{x^{3/2}} \), or possibly some constant times \( \frac{1}{x^{3/2}} \). That is, we need to show that for all \( x \geq 1 \) (i.e. on the domain of integration)
  \[
  \frac{\sqrt{x}}{x^2+x} \leq \frac{A}{x^{3/2}}
  \]
  for some constant \( A \). Let’s try this.
- Since \( x \geq 1 \) we know that \( x^2 + x > x^2 \)
  
  Now take the reciprocal of both sides:
  \[
  \frac{1}{x^2 + x} < \frac{1}{x^2}
  \]

\(^91\) This takes practice, practice and more practice. At the risk of alliteration — please perform plenty of practice problems.
Multiply both sides by $\sqrt{x}$ (which is always positive, so the sign of the inequality does not change)

$$\frac{\sqrt{x}}{x^2 + x} < \frac{\sqrt{x}}{x^2} = \frac{1}{x^{3/2}}$$

So Theorem 1.12.17(a) and Example 1.12.8, with $p = \frac{3}{2}$ do indeed show that the integral $\int_1^\infty \frac{\sqrt{x}}{x^2 + x} \, dx$ converges.

Notice that in this last example we managed to show that the integral exists by finding an integrand that behaved the same way for large $x$. Our intuition then had to be bolstered with some careful inequalities to apply the comparison Theorem 1.12.17. It would be nice to avoid this last step and be able jump from the intuition to the conclusion without messing around with inequalities. Thankfully there is a variant of Theorem 1.12.17 that is often easier to apply and that also fits well with the sort of intuition that we developed to solve Example 1.12.21.

A key phrase in the previous paragraph is “behaves the same way for large $x$”. A good way to formalise this expression — “$f(x)$ behaves like $g(x)$ for large $x$” — is to require that the limit

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

exists and is a finite nonzero number.

Suppose that this is the case and call the limit $L \neq 0$. Then

- the ratio $\frac{f(x)}{g(x)}$ must approach $L$ as $x$ tends to $+\infty$.

- So when $x$ is very large — say $x > B$, for some big number $B$ — we must have that

$$\frac{1}{2}L \leq \frac{f(x)}{g(x)} \leq 2L$$

for all $x > B$

Equivalently, $f(x)$ lies between $\frac{1}{2}g(x)$ and $2Lg(x)$, for all $x \geq B$.

- Consequently, the integral of $f(x)$ converges if and only if the integral of $g(x)$ converges, by Theorems 1.12.17 and 1.12.20.

These considerations lead to the following variant of Theorem 1.12.17.
Theorem 1.12.22 (Limiting comparison).

Let $-\infty < a < \infty$. Let $f$ and $g$ be functions that are defined and continuous for all $x \geq a$ and assume that $g(x) \geq 0$ for all $x \geq a$.

(a) If $\int_a^\infty g(x) \, dx$ converges and the limit

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

exists, then $\int_a^\infty f(x) \, dx$ converges.

(b) If $\int_a^\infty g(x) \, dx$ diverges and the limit

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

exists and is nonzero, then $\int_a^\infty f(x) \, dx$ diverges.

Note that in (b) the limit must exist and be nonzero, while in (a) we only require that the limit exists (it can be zero).

Here is an example of how Theorem 1.12.22 is used.

Example 1.12.23 \( \int_1^\infty \frac{x + \sin x}{e^{-x} + x^2} \, dx \)

Does the integral $\int_1^\infty \frac{x + \sin x}{e^{-x} + x^2} \, dx$ converge or diverge?

Solution.

- Our first task is to identify the potential sources of impropriety for this integral.

- The domain of integration extends to $+\infty$. On the domain of integration the denominator is never zero so the integrand is continuous. Thus the only problem is at $+\infty$.

- Our second task is to develop some intuition about the behavior of the integrand for very large $x$. A good way to start is to think about the size of each term when $x$ becomes big.

- When $x$ is very large:
  - $e^{-x} \ll x^2$, so that the denominator $e^{-x} + x^2 \approx x^2$, and
  - $|\sin x| \leq 1 \ll x$, so that the numerator $x + \sin x \approx x$, and
  - the integrand $\frac{x + \sin x}{e^{-x} + x^2} \approx \frac{x}{x^2} = \frac{1}{x}$.

Notice that we are using $A \ll B$ to mean that “$A$ is much much smaller than $B$”. Similarly $A \gg B$ means “$A$ is much much bigger than $B$”. We don’t really need to be too precise about its meaning beyond this in the present context.
• Now, since \( \int_1^\infty \frac{dx}{x} \) diverges, we would expect \( \int_1^\infty \frac{x + \sin x}{e^{-x} + x^2} \, dx \) to diverge too.

• Our final task is to verify that our intuition is correct. To do so, we set

\[
 f(x) = \frac{x + \sin x}{e^{-x} + x^2} \quad \quad g(x) = \frac{1}{x}
\]

and compute

\[
 \lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{x + \sin x}{e^{-x} + x^2} \cdot \frac{1}{x} \\
= \lim_{x\to\infty} \frac{(1 + \sin x/x)x}{(e^{-x}/x^2 + 1)x^2} \times x \\
= \lim_{x\to\infty} \frac{1 + \sin x/x}{e^{-x}/x^2 + 1} \\
= 1
\]

• Since \( \int_1^\infty g(x) \, dx = \int_1^\infty \frac{dx}{x} \) diverges, by Example 1.12.8 with \( p = 1 \), Theorem 1.12.22(b) now tells us that \( \int_1^\infty f(x) \, dx = \int_1^\infty \frac{x + \sin x}{e^{-x} + x^2} \, dx \) diverges too.
In the previous chapter we defined the definite integral, based on its interpretation as the area of a region in the \(xy\)-plane. We also developed a bunch of theory to help us work with integrals. This abstract definition, and the associated theory, turns out to be extremely useful simply because "areas of regions in the \(xy\)-plane" appear in a huge number of different settings, many of which seem superficially not to involve "areas of regions in the \(xy\)-plane". Here are some examples.

- The work involved in moving a particle or in pumping a fluid out of a reservoir. See section 2.1.
- The average value of a function. See section 2.2.
- The center of mass of an object. See section 2.3.
- The time dependence of temperature. See section 2.4.
- Radiocarbon dating. See section 2.4.

Let us start with the first of these examples.

## 2.1 Work

While computing areas and volumes are nice mathematical applications of integration we can also use integration to compute quantities of importance in physics and statistics. One such quantity is work. Work is a way of quantifying the amount of energy that is required to act against a force. In SI metric units the force \(F\) has units newtons (which

---

1 For example — if your expensive closed-source textbook has fallen on the floor, work quantifies the amount of energy required to lift the object from the floor acting against the force of gravity.

2 SI is short for "le système international d’unités" which is French for "the international system of units". It is the most recent internationally sanctioned version of the metric system, published in 1960. It aims to establish sensible units of measurement (no cubic furlongs per hogshead-Fahrenheit). It defines seven base units — metre (length), kilogram (mass), second (time), kelvin (temperature), ampere (electric current), mole (quantity of substance) and candela (luminous intensity). From these one can then establish derived units — such as metres per second for velocity and speed.
are kilograms meters per second squared), \( x \) has units meters and the work \( W \) has units joules (which are newton meters or kilograms meters squared per second squared).

**Definition 2.1.1.**

The work done by a force \( F(x) \) in moving an object from \( x = a \) to \( x = b \) is

\[
W = \int_{a}^{b} F(x) \, dx
\]

In particular, if the force is a constant, \( F \), independent of \( x \), the work is \( F \cdot (b - a) \).

Here is some motivation for this definition. Consider a particle of mass \( m \) moving along the \( x \)-axis. Let the position of the particle at time \( t \) be \( x(t) \). The particle starts at position \( a \) at time \( \alpha \), moves to the right, finishing at position \( b > a \) at time \( \beta \). While the particle moves, it is subject to a position-dependent force \( F(x) \). Then Newton’s law of motion

\[ m \frac{d^2 x}{dt^2} (t) = F(x(t)) \]

Now consider our definition of work above. It tells us that the work done in moving the particle from \( x = a \) to \( x = b \) is

\[
W = \int_{a}^{b} F(x) \, dx
\]

However, we know the position as a function of time, so we can substitute \( x = x(t) \), \( dx = \frac{dx}{dt} \, dt \) (using Theorem 1.4.6) and rewrite the above integral:

\[
W = \int_{a}^{b} F(x) \, dx = \int_{t=\alpha}^{t=\beta} F(x(t)) \frac{dx}{dt} \, dt
\]

Using Newton’s second law we can rewrite our integrand:

\[
= m \int_{\alpha}^{\beta} \frac{d^2 x}{dt^2} \, dx \, dt
= m \int_{\alpha}^{\beta} \frac{dv}{dt} \, v(t) \, dt
= m \int_{\alpha}^{\beta} \frac{d}{dt} \left( \frac{1}{2} v(t)^2 \right) \, dt
\]

---

3 Specifically, the second of Newton’s three law of motion. These were first published in 1687 in his “Philosophiae Naturalis Principia Mathematica”.

4 It actually says something more graceful in latin - Mutationem motus proportionalem esse vi motrici impressae, et fieri secundum lineam rectam qua vis illa imprimitur. Or — The alteration of motion is ever proportional to the motive force impressed; and is made in the line in which that force is impressed. It is amazing what you can find on the internet.
Applications of Integration

2.1 Work

What happened here? By the chain rule, for any function $f(t)$:

$$\frac{d}{dt} \left( \frac{1}{2} f(t)^2 \right) = f(t)f'(t).$$

In the above computation we have used this fact with $f(t) = v(t)$. Now using the fundamental theorem of calculus (Theorem 1.3.1 part 2), we have

$$W = m \int_{\alpha}^{\beta} \frac{d}{dt} \left( \frac{1}{2} v(t)^2 \right) dt = \frac{1}{2} m v(\beta)^2 - \frac{1}{2} m v(\alpha)^2.$$  

By definition, the function $\frac{1}{2}mv(t)^2$ is the kinetic energy\(^5\) of the particle at time $t$. So the work $W$ of Definition 2.1.1 is the change in kinetic energy from the time the particle was at $x = a$ to the time it was at $x = b$.

Example 2.1.2 (Hooke’s Law)

Imagine that a spring lies along the $x$-axis. The left hand end is fixed to a wall, but the right hand end lies freely at $x = 0$. So the spring is at its “natural length”.

- Now suppose that we wish to stretch out the spring so that its right hand end is at $x = L$.
- Hooke’s Law\(^6\) says that when a (linear) spring is stretched (or compressed) by $x$ units beyond its natural length, it exerts a force of $kx$, where the constant $k$ is the spring constant of that spring.
- In our case, once we have stretched the spring by $x$ units to the right, the spring will be trying to pull back the right hand end by applying a force of magnitude $kx$ directed to the left.
- For us to continue stretching the spring we will have to apply a compensating force of magnitude $kx$ directed to the right. That is, we have to apply the force $F(x) = +kx$.

---

5 This is not a physics text so we will not be too precise. Roughly speaking, kinetic energy is the energy an object possesses due to it being in motion, as opposed to potential energy, which is the energy of the object due to its position in a force field. Leibniz and Bernoulli determined that kinetic energy is proportional to the square of the velocity, while the modern term “kinetic energy” was first used by Lord Kelvin (back while he was still William Thompson).

6 Robert Hooke (1635–1703) was an English contemporary of Isaac Newton (1643–1727). It was in a 1676 letter to Hooke that Newton wrote “If I have seen further it is by standing on the shoulders of Giants.” There is some thought that this was sarcasm and Newton was actually making fun of Hooke, who had a spinal deformity. However at that time Hooke and Newton were still friends. Several years later they did have a somewhat public falling-out over some of Newton’s work on optics.
• So to stretch a spring by $L$ units from its natural length we have to supply the work

$$W = \int_0^L kx\,dx = \frac{1}{2}kL^2$$

Example 2.1.2

Example 2.1.3 (Spring)

A spring has a natural length of 0.1m. If a 12N force is needed to keep it stretched to a length of 0.12m, how much work is required to stretch it from 0.12m to 0.15m?

**Solution.** In order to answer this question we will need to determine the spring constant and then integrate the appropriate function.

• Our first task is to determine the spring constant $k$. We are told that when the spring is stretched to a length of 0.12m, i.e. to a length of $0.12 - 0.1 = 0.02$m beyond its natural length, then the spring generates a force of 12N.

• Hooke’s law states that $F = kx$, so

$$12 = k \cdot 0.02 = k \cdot \frac{2}{100} \quad \text{thus} \quad k = 600.$$

• So to stretch the spring

  – from a length of 0.12m, i.e. a length of $x = 0.12 - 0.1 = 0.02$m beyond its natural length,

  – to a length of 0.15m, i.e. a length of $x = 0.15 - 0.1 = 0.05$m beyond its natural length,

  takes work

$$W = \int_{0.02}^{0.05} kx\,dx = \left[ \frac{1}{2}kx^2 \right]_{0.02}^{0.05}$$

$$= 300(0.05^2 - 0.02^2)$$

$$= 0.63\text{J}$$

Example 2.1.3

Example 2.1.4 (Pumping Out a Reservoir)

A cylindrical reservoir of height $h$ and radius $r$ is filled with a fluid of density $\rho$. We would like to know how much work is required to pump all of the fluid out the top of the reservoir.

We could assign units to these measurements — such as metres for the lengths $h$ and $r$, and kilograms per cubic metre for the density $\rho$. 

210
Solution. We are going to tackle this problem by applying the standard integral calculus “slice into small pieces” strategy. This is how we computed areas and volumes — slice the problem into small pieces, work out how much each piece contributes, and then add up the contributions using an integral.

- Start by slicing the reservoir (or rather the fluid inside it) into thin, horizontal, cylindrical pancakes, as in the figure above. We proceed by determining how much work is required to pump out this pancake volume of fluid.

- Each pancake is a squat cylinder with thickness $dx$ and circular cross section of radius $r$ and area $\pi r^2$. Hence it has volume $\pi r^2 dx$ and mass $\rho \pi r^2 dx$.

- Near the surface of the Earth gravity exerts a downward force of $mg$ on a body of mass $m$. The constant $g = 9.8\text{m/sec}^2$ is called the standard acceleration due to gravity. For us to raise the pancake we have to apply a compensating upward force of $mg$, which, for our pancake, is

$$F = g\rho \times \pi r^2 dx$$

- To remove the pancake at height $x$ from the reservoir we need to raise it to height $h$. So we have to lift it a distance $h - x$ using the force $F = \pi \rho gr^2 dx$, which takes work $\pi \rho gr^2 (h - x) dx$.

- The total work to empty the whole reservoir is

$$W = \int_0^h \pi \rho gr^2 (h - x) dx = \pi \rho g r^2 \int_0^h (h - x) dx$$

$$= \pi \rho g r^2 \left[ hx - \frac{x^2}{2} \right]_0^h$$

$$= \frac{\pi}{2} \rho g r^2 h^2$$

- If we measure lengths in metres and mass in kilograms, then this quantity has units of Joules. If we instead used feet and pounds then this would have units of “foot–pounds”. One foot-pound is equal to 1.355817... Joules.
Example 2.1.5 (Escape Velocity)

Suppose that you shoot a probe straight up from the surface of the Earth — at what initial speed must the probe move in order to escape Earth’s gravity?

Solution. We determine this by computing how much work must be done in order to escape Earth’s gravity. If we assume that all of this work comes from the probe’s initial kinetic energy, then we can establish the minimum initial velocity required.

- The work done in moving a mass from the surface of the Earth to a height \( h \) above the surface is

\[
W = \int_0^h F(x)\,dx
\]

where \( F(x) \) is the gravitational force acting on the mass at height \( x \) above the Earth’s surface.

- The gravitational force\(^{11}\) of the Earth acting on a particle of mass \( m \) at a height \( x \) above the surface of the Earth is

\[
F = -\frac{GMm}{(R + x)^2},
\]

where \( G \) is the gravitational constant, \( M \) is the mass of the Earth and \( R \) is the radius of the Earth. Note that \( R + x \) is the distance from the object to the centre of the Earth. Additionally, note that this force is negative because gravity acts downward.

- So the work done by gravity on the probe, as it travels from the surface of the Earth to a height \( h \), is

\[
W = -\int_0^h \frac{GMm}{(R + x)^2}\,dx
= -GMm\int_0^h \frac{1}{(R + x)^2}\,dx
\]

\(^{11}\) Newton published his inverse square law of universal gravitation in his Principia in 1687. His law states that the gravitational force between two masses \( m_1 \) and \( m_2 \) is

\[
F = -G\frac{m_1m_2}{r^2}
\]

where \( r \) is the distance separating the (centres of the) masses and \( G = 6.674 \times 10^{11}\text{Nm}^2/\text{kg}^2 \) is the gravitational constant. Notice that \( r \) measures the separation between the centres of the masses not the distance between the surfaces of the objects.

Also, do not confuse \( G \) with \( g \) — standard acceleration due to gravity. The first measurement of \( G \) was performed by Henry Cavendish in 1798 — the interested reader should look up the “Cavendish experiment” for details of this very impressive work.
A quick application of the substitution rule with $u = R + x$ gives

$$\begin{align*}
&= -GMm \int_{u(0)}^{u(h)} \frac{1}{u^2} \, du \\
&= -GMm \left[ \frac{-1}{u} \right]_{u=R}^{u=R+h} \\
&= \frac{GMm}{R+h} - \frac{GMm}{R}
\end{align*}$$

- So if the probe completely escapes the Earth and travels all the way to $h = \infty$, gravity does work

$$\lim_{h \to \infty} \left[ \frac{GMm}{R+h} - \frac{GMm}{R} \right] = -\frac{GMm}{R}$$

The minus sign means that gravity has removed energy $\frac{GMm}{R}$ from the probe.

- To finish the problem we need one more assumption. Let us assume that all of this energy comes from the probe’s initial kinetic energy and that the probe is not fitted with any sort of rocket engine. Hence the initial kinetic energy $\frac{1}{2}mv^2$ (coming from an initial velocity $v$) must be at least as large as the work computed above. That is we need

$$\frac{1}{2}mv^2 \geq \frac{GMm}{R}$$

which rearranges to give

$$v \geq \sqrt{\frac{2GM}{R}}$$

- The right hand side of this inequality, $\sqrt{\frac{2GM}{R}}$, is called the escape velocity.

---

### 2.2 Averages

Another frequent application of integration is computing averages and other statistical quantities. We will not spend too much time on this topic — that is best left to a proper course in statistics — however, we will demonstrate the application of integration to the problem of computing averages.

---

12 Awful pun. The two main approaches to statistics are frequentism and Bayesianism; the latter named after Bayes’ Theorem which is, in turn, named for Reverand Thomas Bayes. While this (both the approaches to statistics and their history and naming) is a very interesting and quite philosophical topic, it is beyond the scope of this course. The interested reader has plenty of interesting reading here to interest them.
Let us start with the definition of the average of a finite set of numbers.

**Definition 2.2.1.**

The average (mean) of a set of \( n \) numbers \( y_1, y_2, \ldots, y_n \) is

\[
y_{\text{ave}} = \bar{y} = \langle y \rangle = \frac{y_1 + y_2 + \cdots + y_n}{n}
\]

The notations \( y_{\text{ave}} \), \( \bar{y} \) and \( \langle y \rangle \) are all commonly used to represent the average.

Now suppose that we want to take the average of a function \( f(x) \) with \( x \) running continuously from \( a \) to \( b \). How do we even define what that means? A natural approach is to

- select, for each natural number \( n \), a sample of \( n \) more or less uniformly distributed values of \( x \) between \( a \) and \( b \),
- take the average of the values of \( f \) at the selected points,
- and then take the limit as \( n \) tends to infinity.

Unsurprisingly, this process looks very much like how we computed areas and volumes previously. So let’s get to it.

- First fix any natural number \( n \).
- Subdivide the interval \( a \leq x \leq b \) into \( n \) equal subintervals, each of width \( \Delta x = \frac{b-a}{n} \).
- The subinterval number \( i \) runs from \( x_{i-1} \) to \( x_i \) with \( x_i = a + \frac{i(b-a)}{n} \).
- Select, for each \( 1 \leq i \leq n \), one value of \( x \) from subinterval number \( i \) and call it \( x_i^* \). So \( x_{i-1} \leq x_i^* \leq x_i \).

---

13 We are being a little loose here with the distinction between mean and average. To be much more pedantic — the average is the arithmetic mean. Other interesting “means” are the geometric and harmonic means:

- arithmetic mean \( = \frac{1}{n} (y_1 + y_2 + \cdots + y_n) \)
- geometric mean \( = (y_1 \cdot y_2 \cdots y_n)^{1/n} \)
- harmonic mean \( = \left[ \frac{1}{n} \left( \frac{1}{y_1} + \frac{1}{y_2} + \cdots + \frac{1}{y_n} \right) \right]^{-1} \)

All of these quantities, along with the median and mode, are ways to measure the typical value of a set of numbers. They all have advantages and disadvantages — another interesting topic beyond the scope of this course, but plenty of fodder for the interested reader and their favourite search engine. But let us put pedantry (and beyond-the-scope-of-the-course-reading) aside and just use the terms average and mean interchangably for our purposes here.
• The average value of $f$ at the selected points is
\[
\frac{1}{n} \sum_{i=1}^{n} f(x_i) = \frac{1}{b-a} \sum_{i=1}^{n} f(x_i^*) \Delta x
\]
since $\Delta x = \frac{b-a}{n}$
giving us a Riemann sum.

Now when we take the limit $n \to \infty$ we get exactly $\frac{1}{b-a} \int_{a}^{b} f(x) \, dx$. That’s why we define

**Definition 2.2.2.**

Let $f(x)$ be an integrable function defined on the interval $a \leq x \leq b$. The average value of $f$ on that interval is

\[
f_{\text{ave}} = \bar{f} = \langle f \rangle = \frac{1}{b-a} \int_{a}^{b} f(x) \, dx
\]

Consider the case when $f(x)$ is positive. Then rewriting Definition 2.2.2 as

\[
f_{\text{ave}} (b-a) = \int_{a}^{b} f(x) \, dx
\]
gives us a link between the average value and the area under the curve. The right-hand side is the area of the region

\[
\{(x,y) \mid a \leq x \leq b, \ 0 \leq y \leq f(x) \}
\]
while the left-hand side can be seen as the area of a rectangle of width $b-a$ and height $f_{\text{ave}}$. Since these areas must be the same, we interpret $f_{\text{ave}}$ as the height of the rectangle which has the same width and the same area as $\{(x,y) \mid a \leq x \leq b, \ 0 \leq y \leq f(x) \}$.

Let us start with a couple of simple examples and then work our way up to harder ones.

**Example 2.2.3**

Let $f(x) = x$ and $g(x) = x^2$ and compute their average values over $1 \leq x \leq 5$.

**Solution.** We can just plug things into the definition.

\[
f_{\text{ave}} = \frac{1}{5-1} \int_{1}^{5} x \, dx = \frac{1}{4} \left[ \frac{x^2}{2} \right]_{1}^{5} = \frac{1}{8} (25 - 1) = \frac{24}{8} = 3
\]
as we might expect. And then

\[ g_{\text{ave}} = \frac{1}{5 - 1} \int_{1}^{5} x^2 \, dx \]

\[ = \frac{1}{4} \left[ \frac{x^3}{3} \right]_{1}^{5} \]

\[ = \frac{1}{12} (125 - 1) = \frac{124}{12} = 10 \frac{1}{3} \]

Example 2.2.3

Something a little more trigonometric

Example 2.2.4

Find the average value of \( \sin(x) \) over \( 0 \leq x \leq \pi/2 \).

Solution. Again, we just need the definition.

\[ \text{average} = \frac{1}{\pi/2 - 0} \int_{0}^{\pi/2} \sin(x) \, dx \]

\[ = \frac{2}{\pi} \left[ -\cos(x) \right]_{0}^{\pi/2} \]

\[ = \frac{2}{\pi} (-\cos(\pi/2) + \cos(0)) \]

\[ = \frac{2}{\pi} \]

Example 2.2.4

We could keep going... But better to do some more substantial examples.

Example 2.2.5 (Average velocity)

Let \( x(t) \) be the position at time \( t \) of a car moving along the \( x \)-axis. The velocity of the car at time \( t \) is the derivative \( v(t) = x'(t) \). The average velocity of the car over the time interval \( a \leq t \leq b \) is

\[ v_{\text{ave}} = \frac{1}{b - a} \int_{a}^{b} v(t) \, dt \]

\[ = \frac{1}{b - a} \int_{a}^{b} x'(t) \, dt \]

\[ = \frac{x(b) - x(a)}{b - a} \quad \text{by the fundamental theorem of calculus.} \]
The numerator in this formula is just the displacement (net distance travelled — if \( x'(t) \geq 0 \), it’s the distance travelled) between time \( a \) and time \( b \) and the denominator is just the time it took.

Notice that this is exactly the formula we used way back at the start of your differential calculus class to help introduce the idea of the derivative. Of course this is a very circuitous way to get to this formula — but it is reassuring that we get the same answer.

A very physics example.

**Example 2.2.6 (Peak vs RMS voltage)**

When you plug a light bulb into a socket\(^{14}\) and turn it on, it is subjected to a voltage

\[
V(t) = V_0 \sin(\omega t - \delta)
\]

where

- \( V_0 = 170 \) volts,
- \( \omega = 2\pi \times 60 \) (which corresponds to 60 cycles per second\(^{15}\)) and
- the constant \( \delta \) is an (unimportant) phase. It just shifts the time at which the voltage is zero

The voltage \( V_0 \) is the “peak voltage” — the maximum value the voltage takes over time. More typically we quote the “root mean square” voltage\(^{16}\) (or RMS-voltage). In this example we explain the difference, but to simplify the calculations, let us simplify the voltage function and just use

\[
V(t) = V_0 \sin(t)
\]

Since the voltage is a sine-function, it takes both positive and negative values. If we

\(^{14}\) A normal household socket delivers alternating current, rather than the direct current USB supplies. At the risk of yet another “the interested reader” suggestion — the how and why household plugs supply AC current is another worthwhile and interesting digression from studying integration. The interested reader should look up the “War of Currents”. The diligent and interested reader should bookmark this, finish the section and come back to it later.

\(^{15}\) Some countries supply power at 50 cycles per second. Japan actually supplies both — 50 cycles in the east of the country and 60 in the west.

\(^{16}\) This example was written in North America where the standard voltage supplied to homes is 120 volts. Most of the rest of the world supplies homes with 240 volts. The main reason for this difference is the development of the light bulb. The USA electrified earlier when the best voltage for bulb technology was 110 volts. As time went on, bulb technology improved and countries that electrified later took advantage of this (and the cheaper transmission costs that come with higher voltage) and standardised at 240 volts. So many digressions in this section!
take its simple average over 1 period then we get

\[ V_{\text{ave}} = \frac{1}{2\pi} \int_0^{2\pi} V_0 \sin(t) \, dt \]

\[ = \frac{V_0}{2\pi} \left[ -\cos(t) \right]_0^{2\pi} \]

\[ = \frac{V_0}{2\pi} (-\cos(2\pi) + \cos(0)) = \frac{V_0}{2\pi} (-1 + 1) \]

\[ = 0 \]

This is clearly not a good indication of the typical voltage.

What we actually want here is a measure of how far the voltage is from zero. Now we could do this by taking the average of \(|V(t)|\), but this is a little harder to work with. Instead we take the average of the square\(^{17}\) of the voltage (so it is always positive) and then take the square root at the end. That is

\[
V_{\text{rms}} = \sqrt{\frac{1}{2\pi} \int_0^{2\pi} V(t)^2 \, dt}
\]

\[ = \sqrt{\frac{1}{2\pi} \int_0^{2\pi} V_0^2 \sin^2(t) \, dt} \]

\[ = \sqrt{\frac{V_0^2}{2\pi} \int_0^{2\pi} \sin^2(t) \, dt} \]

This is called the “root mean square” voltage.

Though we do know how to integrate sine and cosine, we don’t (yet) know how to integrate their squares. A quick look at double-angle formulas\(^{18}\) gives us a way to eliminate the square:

\[ \cos(2\theta) = 1 - 2\sin^2 \theta \implies \sin^2 \theta = \frac{1 - \cos(2\theta)}{2} \]

---

\(^{17}\) For a finite set of numbers one can compute the “quadratic mean” which is another way to generalise the notion of the average:

\[ \text{quadratic mean} = \sqrt{\frac{1}{n} (y_1^2 + y_2^2 + \cdots + y_n^2)} \]

\(^{18}\) A quick glance at Appendix A.14 will refresh your memory.
Applications of Integration

2.2 Averages

Using this we manipulate our integrand a little more:

\[
V_{\text{rms}} = \sqrt{\frac{V_0^2}{2\pi} \int_0^{2\pi} \frac{1}{2} (1 - \cos(2t)) \, dt}
\]

\[
= \sqrt{\frac{V_0^2}{4\pi} \left[ t - \frac{1}{2} \sin(2t) \right]_0^{2\pi}}
\]

\[
= \sqrt{\frac{V_0^2}{4\pi} \left( 2\pi - \frac{1}{2} \sin(4\pi) - 0 + \frac{1}{2} \sin(0) \right)}
\]

\[
= \sqrt{\frac{V_0^2}{4\pi} \cdot 2\pi}
\]

\[
= \frac{V_0}{\sqrt{2}}
\]

So if the peak voltage is 170 volts then the RMS voltage is \( \frac{170}{\sqrt{2}} \approx 120.2 \).

Example 2.2.6

Continuing this very physics example:

Let us take our same light bulb with voltage (after it is plugged in) given by

\[ V(t) = V_0 \sin(\omega t - \delta) \]

where

- \( V_0 \) is the peak voltage,
- \( \omega = 2\pi \times 60 \), and
- the constant \( \delta \) is an (unimportant) phase.

If the light bulb is “100 watts”, then what is its resistance?

To answer this question we need the following facts from physics.

- If the light bulb has resistance \( R \) ohms, this causes, by Ohm’s law, a current of

\[ I(t) = \frac{1}{R} V(t) \]

(amps) to flow through the light bulb.

- The current \( I \) is the number of units of charge moving through the bulb per unit time.

- The voltage is the energy required to move one unit of charge through the bulb.

- The power is the energy used by the bulb per unit time and is measured in watts.

219
So the power is the product of the current times the voltage and, so

\[ P(t) = I(t)V(t) = \frac{V(t)^2}{R} = \frac{V_0^2}{R} \sin^2(\omega t - \delta) \]

The average power used over the time interval \(a \leq t \leq b\) is

\[ P_{\text{ave}} = \frac{1}{b-a} \int_a^b P(t) \, dt = \frac{V_0^2}{R(b-a)} \int_a^b \sin^2(\omega t - \delta) \, dt \]

Notice that this is almost exactly the form we had in the previous example when computing the root mean square voltage.

Again we simplify the integrand using the identity

\[ \cos(2\theta) = 1 - 2\sin^2 \theta \implies \sin^2 \theta = \frac{1 - \cos(2\theta)}{2} \]

So

\[ P_{\text{ave}} = \frac{1}{b-a} \int_a^b P(t) \, dt = \frac{V_0^2}{2R(b-a)} \int_a^b \left[ 1 - \cos(2\omega t - 2\delta) \right] \, dt \]

\[ = \frac{V_0^2}{2R(b-a)} \left[ t - \frac{\sin(2\omega t - 2\delta)}{2\omega} \right]_a^b \]

\[ = \frac{V_0^2}{2R(b-a)} \left[ b - a - \frac{\sin(2\omega b - 2\delta)}{2\omega} + \frac{\sin(2\omega a - 2\delta)}{2\omega} \right] \]

\[ = \frac{V_0^2}{2R} - \frac{V_0^2}{4\omega R(b-a)} \left[ \sin(2\omega b - 2\delta) - \sin(2\omega a - 2\delta) \right] \]

In the limit as the length of the time interval \(b - a\) tends to infinity, this converges to \(\frac{V_0^2}{2R}\). The resistance \(R\) of a "100 watt bulb" obeys

\[ \frac{V_0^2}{2R} = 100 \quad \text{so that} \quad R = \frac{V_0^2}{200}. \]

We finish this example off with two side remarks.

- If we translate the peak voltage to the root mean square voltage using
  \[ V_0 = V_{\text{rms}} \cdot \sqrt{2} \]
  then we have
  \[ P = \frac{V_{\text{rms}}^2}{R} \]

- If we were using direct voltage rather than alternating current then the computation is much simpler. The voltage and current are constants, so
  \[ P = V \cdot I \quad \text{but} \quad I = V / R \text{ by Ohm's law} \]
  \[ = \frac{V^2}{R} \]

So if we have a direct current giving voltage equal to the root mean square voltage, then we would expend the same power.
Applications of Integration

2.3 Centre of Mass and Torque

Optional — Return to the mean value theorem

One last application of Definition 2.2.2. The following theorem can be thought of as an analogue of the mean-value theorem (which was covered in your differential calculus class) but for integrals. The theorem says that a continuous function has to take its average value exactly somewhere. For example, if you went for a drive along the x-axis and you were at \( x(a) \) at time \( a \) and at \( x(b) \) at time \( b \), then your velocity \( x'(t) \) had to be exactly your average velocity \( \frac{x(b) - x(a)}{b - a} \) at some time \( t \) between \( a \) and \( b \). In particular, if your average velocity was greater than the speed limit, you were definitely speeding at some point during the trip. This is, of course, no great surprise.

**Theorem 2.2.8 (Mean Value Theorem for Integrals).**

Let \( f(x) \) be a continuous function on the interval \( a \leq x \leq b \). Then there is some \( c \) obeying \( a \leq c \leq b \) such that

\[
\frac{1}{b - a} \int_a^b f(x) \, dx = f(c) \quad \text{or} \quad \int_a^b f(x) \, dx = f(c) \, (b - a)
\]

Here is why the theorem is true. Let \( M \) and \( m \) be the largest and smallest values, respectively, that \( f(x) \) takes for \( x \) between \( a \) and \( b \). Then \( m \leq f(x) \leq M \) for all \( a \leq x \leq b \), so that

\[
m(b - a) \leq \int_a^b f(x) \, dx \leq M(b - a) \quad \iff \quad m \leq \frac{1}{b - a} \int_a^b f(x) \, dx \leq M
\]

As \( x \) runs from \( a \) to \( b \) the continuous function \( f(x) \) takes all values between \( m \) and \( M \) including, in particular, \( \frac{1}{b - a} \int_a^b f(x) \).

2.3 Centre of Mass and Torque

2.3.1 Centre of Mass

If you support a body at its center of mass (in a uniform gravitational field) it balances perfectly. That’s the definition of the center of mass of the body. If the body consists of a
finite number of masses $m_1, \ldots, m_n$ attached to an infinitely strong, weightless (idealized) rod with mass number $i$ attached at position $x_i$, then the center of mass is at the (weighted) average value of $x$:

$$\bar{x} = \frac{\sum_{i=1}^{n} m_i x_i}{\sum_{i=1}^{n} m_i}$$

The denominator $m = \sum_{i=1}^{n} m_i$ is the total mass of the body. This formula for the center of mass is derived in the following (optional) section. See (2.3.6).

For many (but certainly not all) purposes an (extended rigid) body acts like a point particle located at its center of mass. For example it is very common to treat the Earth as a point particle. Here is a more detailed example in which we think of a body as being made up of a number of component parts and compute the center of mass of the body as a whole by using the center of masses of the component parts. Suppose that we have a dumbbell which consists of

- a left end made up of particles of masses $m_{l,1}, \ldots, m_{l,3}$ located at $x_{l,1}, \ldots, x_{l,3}$ and
- a right end made up of particles of masses $m_{r,1}, \ldots, m_{r,4}$ located at $x_{r,1}, \ldots, x_{r,4}$ and
- an infinitely strong, weightless (idealized) rod joining all of the particles.

Then the mass and center of mass of the left end are

$$M_l = m_{l,1} + \cdots + m_{l,3} \quad \bar{x}_l = \frac{m_{l,1} x_{l,1} + \cdots + m_{l,3} x_{l,3}}{M_l}$$

and the mass and center of mass of the right end are

$$M_r = m_{r,1} + \cdots + m_{r,4} \quad \bar{x}_r = \frac{m_{r,1} x_{r,1} + \cdots + m_{r,4} x_{r,4}}{M_r}$$

The mass and center of mass of the entire dumbbell are

$$M = m_{l,1} + \cdots + m_{l,3} + m_{r,1} + \cdots + m_{r,4} = M_l + M_r$$

$$\bar{x} = \frac{m_{l,1} x_{l,1} + \cdots + m_{l,3} x_{l,3} + m_{r,1} x_{r,1} + \cdots + m_{r,4} x_{r,4}}{M} = \frac{M_l \bar{x}_l + M_r \bar{x}_r}{M_l + M_r}$$

So we can compute the center of mass of the entire dumbbell by treating it as being made up of two point particles, one of mass $M_l$ located at the centre of mass of the left end, and one of mass $M_r$ located at the center of mass of the right end.

Now we’ll extend the above ideas to cover more general classes of bodies. If the body consists of mass distributed continuously along a straight line, say with mass density $\rho(x) \text{kg/m}$ and with $x$ running from $a$ to $b$, rather than consisting of a finite number of point masses, the formula for the center of mass becomes

$$\bar{x} = \frac{\int_a^b x \rho(x) \, dx}{\int_a^b \rho(x) \, dx}$$

Think of $\rho(x) \, dx$ as the mass of the “almost point particle” between $x$ and $x + dx$. 

---

222
If the body is a two dimensional object, like a metal plate, lying in the $xy$–plane, its center of mass is a point $(\bar{x}, \bar{y})$ with $\bar{x}$ being the (weighted) average value of the $x$–coordinate over the body and $\bar{y}$ being the (weighted) average value of the $y$–coordinate over the body. To be concrete, suppose the body fills the region

$$\{ (x, y) \mid a \leq x \leq b, B(x) \leq y \leq T(x) \}$$

in the $xy$–plane. For simplicity, we will assume that the density of the body is a constant, say $\rho$. When the density is constant, the center of mass is also called the centroid and is thought of as the geometric center of the body.

To find the centroid of the body, we use the use our standard “slicing” strategy. We slice the body into thin vertical strips, as illustrated in the figure below. Here is a detailed description of a generic strip.

- The strip has width $dx$.
- Each point of the strip has essentially the same $x$–coordinate. Call it $x$.
- The top of the strip is at $y = T(x)$ and the bottom of the strip is at $y = B(x)$.
- So the strip has
  - height $T(x) - B(x)$
  - area $[T(x) - B(x)] \, dx$
  - mass $\rho [T(x) - B(x)] \, dx$
  - centroid, i.e. middle point, $(x, \frac{B(x) + T(x)}{2})$.

In computing the centroid of the entire body, we may treat each strip as a single particle of mass $\rho [T(x) - B(x)] \, dx$ located at $(x, \frac{B(x) + T(x)}{2})$. So the mass of the entire body is

$$M = \rho \int_{a}^{b} [T(x) - B(x)] \, dx = \rho A$$  \hspace{1cm} (2.3.1a)
where \( A = \int_a^b [T(x) - B(x)] \, dx \) is the area of the region. The coordinates of the centroid are

\[
\bar{x} = \frac{\int_a^b x \, \rho[T(x) - B(x)] \, dx}{M} = \frac{\int_a^b x[T(x) - B(x)] \, dx}{A} \tag{2.3.1b}
\]

\[
\bar{y} = \frac{\int_a^b \frac{B(x) + T(x)}{2} \rho[T(x) - B(x)] \, dx}{M} = \frac{\int_a^b \frac{1}{2} [T(x)^2 - B(x)^2] \, dx}{2A} \tag{2.3.1c}
\]

We can of course also slice up the body using horizontal slices. If the body has constant density \( \rho \) and fills the region

\[
\{(x, y) \mid L(y) \leq x \leq R(y), \ c \leq y \leq d\}
\]

then the same computation as above gives the mass of the body to be

\[
M = \rho \int_c^d [R(y) - L(y)] \, dy = \rho A \tag{2.3.2a}
\]

where \( A = \int_c^d [R(y) - L(y)] \, dy \) is the area of the region, and gives the coordinates of the centroid to be

\[
\bar{x} = \frac{\int_c^d \frac{R(y) + L(y)}{2} \rho[R(y) - L(y)] \, dy}{M} = \frac{\int_c^d \frac{1}{2} [R(y)^2 - L(y)^2] \, dx}{2A} \tag{2.3.2b}
\]

\[
\bar{y} = \frac{\int_c^d y \, \rho[R(y) - L(y)] \, dy}{M} = \frac{\int_c^d y[R(y) - L(y)] \, dy}{A} \tag{2.3.2c}
\]

**Example 2.3.1**

Find the \( x \)-coordinate of the centroid (centre of gravity) of the plane region \( R \) that lies in the first quadrant \( x \geq 0, \ y \geq 0 \) and inside the ellipse \( 4x^2 + 9y^2 = 36 \). (The area bounded by the ellipse \( \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \) is \( \pi ab \) square units.)
Applications of Integration

2.3 Centre of Mass and Torque

Solution. In standard form $4x^2 + 9y^2 = 36$ is $\frac{x^2}{9} + \frac{y^2}{4} = 1$. So, on $R$, $x$ runs from 0 to 3 and $R$ has area $A = \frac{1}{4}\pi \times 3 \times 2 = \frac{3}{2}\pi$. For each fixed $x$, between 0 and 3, $y$ runs from 0 to $2\sqrt{1 - \frac{x^2}{9}}$. So, applying (2.3.1.b) with $a = 0$, $b = 3$, $T(x) = 2\sqrt{1 - \frac{x^2}{9}}$ and $B(x) = 0$,$$
\bar{x} = \frac{1}{A} \int_0^3 x T(x) \, dx = \frac{1}{A} \int_0^3 x 2\sqrt{1 - \frac{x^2}{9}} \, dx = \frac{4}{3\pi} \int_0^3 x \sqrt{1 - \frac{x^2}{9}} \, dx$

Sub in $u = 1 - \frac{x^2}{9}, \, du = -\frac{2}{9}x \, dx$.

$$\bar{x} = \frac{-9}{23\pi} \int_1^0 \sqrt{u} \, du = \frac{-9}{23\pi} \left[ \frac{u^{3/2}}{3/2} \right]_1^0 = \frac{-9}{23\pi} \left[ \frac{2}{3} \right] = \frac{4}{\pi}$$

Example 2.3.2

Find the centroid of the quarter circular disk $x \geq 0, y \geq 0, x^2 + y^2 \leq r^2$.

Solution. By symmetry, $\bar{x} = \bar{y}$. The area of the quarter disk is $A = \frac{1}{4}\pi r^2$. By (2.3.1.b) with $a = 0$, $b = r$, $T(x) = \sqrt{r^2 - x^2}$ and $B(x) = 0$,$$
\bar{x} = \frac{1}{A} \int_0^r x \sqrt{r^2 - x^2} \, dx$$

To evaluate the integral, sub in $u = r^2 - x^2, \, du = -2x \, dx$.

$$\int_0^r x \sqrt{r^2 - x^2} \, dx = \int_0^{r^2} \sqrt{u} \, du = \frac{1}{2} \left[ \frac{u^{3/2}}{3/2} \right]_0^{r^2} = \frac{r^3}{3} \quad (2.3.3)$$
So
\[ \bar{x} = \frac{4}{\pi r^2} \left[ \frac{r^3}{3} \right] = \frac{4r}{3\pi} \]

As we observed above, we should have \( \bar{x} = \bar{y} \). But, just for practice, let's compute \( \bar{y} \) by the integral formula (2.3.1.c), again with \( a = 0 \), \( b = r \), \( T(x) = \sqrt{r^2 - x^2} \) and \( B(x) = 0 \),

\[
\bar{y} = \frac{1}{2A} \int_0^r (\sqrt{r^2 - x^2})^2 \, dx = \frac{2}{\pi r^2} \int_0^r (r^2 - x^2) \, dx
\]

\[ = \frac{2}{\pi r^2} \left[ r^2 x - \frac{x^3}{3} \right]_0^r = \frac{2}{\pi r^2} \frac{2r^3}{3} \]

\[ = \frac{4r}{3\pi} \]

as expected.

Example 2.3.2

Example 2.3.3

Find the centroid of the region \( R \) in the diagram.

Solution. By symmetry, \( \bar{x} = \bar{y} \). The region \( R \) is a \( 2 \times 2 \) square with one quarter of a circle of radius 1 removed and so has area \( 2 \times 2 - \frac{1}{4} \pi = \frac{16 - \pi}{4} \). The top of \( R \) is \( y = T(x) = 2 \). The bottom is \( y = B(x) \) with \( B(x) = \sqrt{1 - x^2} \) when \( 0 \leq x \leq 1 \) and \( B(x) = 0 \) when \( 1 \leq x \leq 2 \). So

\[
\bar{y} = \bar{x} = \frac{1}{A} \left[ \int_0^1 x[2 - \sqrt{1 - x^2}] \, dx + \int_1^2 x[2 - 0] \, dx \right]
\]

\[ = \frac{4}{16 - \pi} \left[ x^2 \bigg|_0^1 + x^2 \bigg|_1^2 - \int_0^1 x\sqrt{1 - x^2} \, dx \right] \]

\[ = \frac{4}{16 - \pi} \left[ 4 - \frac{1}{3} \right] \quad \text{by (2.3.3) with } r = 1 \]

\[ = \frac{44}{48 - 3\pi} \]
Example 2.3.4

Prove that the centroid of any triangle is located at the point of intersection of the medians. A median of a triangle is a line segment joining a vertex to the midpoint of the opposite side.

Solution. Choose a coordinate system so that the vertices of the triangle are located at \((a, 0)\), \((0, b)\) and \((c, 0)\). (In the figure below, \(a\) is negative.) The line joining \((a, 0)\) and \((0, b)\) has equation \(bx + ay = ab\). (Check that \((a, 0)\) and \((0, b)\) both really are on this line.) The line joining \((c, 0)\) and \((0, b)\) has equation \(bx + cy = bc\). (Check that \((c, 0)\) and \((0, b)\) both really are on this line.) Hence for each fixed \(y\) between 0 and \(b\), \(x\) runs from \(a - \frac{a}{b}y\) to \(c - \frac{c}{b}y\).

We’ll use horizontal strips to compute \(\bar{x}\) and \(\bar{y}\). We could just apply (2.3.2) with \(c = 0\), \(d = b\), \(R(y) = \frac{c}{b}(b - y)\) (which is gotten by solving \(bx + cy = bc\) for \(x\)) and \(L(y) = \frac{a}{b}(b - y)\) (which is gotten by solving \(bx + ay = ab\) for \(x\)).

But rather than memorizing or looking up those formulae, we’ll derive them for this example. So consider a thin strip at height \(y\) as illustrated in the figure above.

- The strip has length
  \[
  \ell(y) = \left[ \frac{c}{b}(b - y) - \frac{a}{b}(b - y) \right] = \frac{c - a}{b}(b - y)
  \]
- The strip has width \(dy\).
- On this strip, \(y\) has average value \(y\).
- On this strip, \(x\) has average value \(\frac{1}{2} \left[ \frac{c}{b}(b - y) + \frac{a}{b}(b - y) \right] = \frac{a + c}{2b}(b - y)\).

As the area of the triangle is \(A = \frac{1}{2}(c - a)b\),

\[
\bar{y} = \frac{1}{A} \int_0^b y \ell(y) \, dy = \frac{2}{(c - a)b} \int_0^b y \frac{c - a}{b}(b - y) \, dy = \frac{2}{b^2} \int_0^b (by - y^2) \, dy = \frac{2}{b^2} \left( \frac{b^2}{2} - \frac{b^3}{3} \right)
\]

\[
= \frac{2b^3}{6} = \frac{b}{3}
\]


2.3.2 • Optional — Torque

Newton’s law of motion says that the position \( x(t) \) of a single particle moving under the influence of a force \( F \) obeys \( mx''(t) = F \). Similarly, the positions \( x_i(t), 1 \leq i \leq n \), of a set of particles moving under the influence of forces \( F_i \) obey \( mx_i''(t) = F_i, 1 \leq i \leq n \). Often systems of interest consist of some small number of rigid bodies. Suppose that we are interested in the motion of a single rigid body, say a piece of wood. The piece of wood is made up of a huge number of atoms. So the system of equations determining the motion of all of the individual atoms in the piece of wood is huge. On the other hand, because the piece of wood is rigid, its configuration is completely determined by the position of, for example, its centre of mass and its orientation. (Rather than get into what is precisely meant by “orientation”, let’s just say that it is certainly determined by, for example, the
positions of a few of the corners of the piece of wood). It is possible to extract from the huge system of equations that determine the motion of all of the individual atoms, a small system of equations that determine the motion of the centre of mass and the orientation. We can avoid some vector analysis, that is beyond the scope of this course, by assuming that our rigid body is moving in two rather than three dimensions.

So, imagine a piece of wood moving in the $xy$–plane. Furthermore, imagine that the piece of wood consists of a huge number of particles joined by a huge number of weightless but very strong steel rods. The steel rod joining particle number one to particle number two just represents a force acting between particles number one and two. Suppose that

- there are $n$ particles, with particle number $i$ having mass $m_i$
- at time $t$, particle number $i$ has $x$–coordinate $x_i(t)$ and $y$–coordinate $y_i(t)$
- at time $t$, the external force (gravity and the like) acting on particle number $i$ has $x$–coordinate $H_i(t)$ and $y$–coordinate $V_i(t)$. Here $H$ stands for horizontal and $V$ stands for vertical.
- at time $t$, the force acting on particle number $i$, due to the steel rod joining particle number $i$ to particle number $j$ has $x$–coordinate $H_{ij}(t)$ and $y$–coordinate $V_{ij}(t)$. If there is no steel rod joining particles number $i$ and $j$, just set $H_{ij}(t) = V_{ij}(t) = 0$. In particular, $H_{ii}(t) = V_{ii}(t) = 0$.

The only assumptions that we shall make about the steel rod forces are

(A1) for each $i \neq j$, $H_{ij}(t) = -H_{ji}(t)$ and $V_{ij}(t) = -V_{ji}(t)$. In words, the steel rod joining particles $i$ and $j$ applies equal and opposite forces to particles $i$ and $j$.

(A2) for each $i \neq j$, there is a function $M_{ij}(t)$ such that $H_{ij}(t) = M_{ij}(t) [x_i(t) - x_j(t)]$ and $V_{ij}(t) = M_{ij}(t) [y_i(t) - y_j(t)]$. In words, the force due to the rod joining particles $i$ and $j$ acts parallel to the line joining particles $i$ and $j$. For (A1) to be true, we need $M_{ij}(t) = M_{ji}(t)$.

Newton’s law of motion, applied to particle number $i$, now tells us that

$$m_i x_i''(t) = H_i(t) + \sum_{j=1}^{n} H_{ij}(t) \quad (X_i)$$

$$m_i y_i''(t) = V_i(t) + \sum_{j=1}^{n} V_{ij}(t) \quad (Y_i)$$
Adding up all of the equations \((X_i)\), for \(i = 1, 2, 3, \ldots, n\) and adding up all of the equations \((Y_i)\), for \(i = 1, 2, 3, \ldots, n\) gives

\[
\sum_{i=1}^{n} m_i x_i''(t) = \sum_{i=1}^{n} H_i(t) + \sum_{1 \leq i, j \leq n} H_{ij}(t) \quad (\Sigma_i X_i)
\]

\[
\sum_{i=1}^{n} m_i y_i''(t) = \sum_{i=1}^{n} V_i(t) + \sum_{1 \leq i, j \leq n} V_{ij}(t) \quad (\Sigma_i Y_i)
\]

The sum \(\sum_{1 \leq i, j \leq n} H_{ij}(t)\) contains \(H_{1,2}(t)\) exactly once and it also contains \(H_{2,1}(t)\) exactly once and these two terms cancel exactly, by assumption (A1). In this way, all terms in \(\sum_{1 \leq i, j \leq n} H_{ij}(t)\) with \(i \neq j\) exactly cancel. All terms with \(i = j\) are assumed to be zero. So \(\sum_{1 \leq i, j \leq n} H_{ij}(t) = 0\). Similarly, \(\sum_{1 \leq i, j \leq n} V_{ij}(t) = 0\), so the equations \((\Sigma_i X_i)\) and \((\Sigma_i Y_i)\) simplify to

\[
\sum_{i=1}^{n} m_i x_i''(t) = \sum_{i=1}^{n} H_i(t) \quad (\Sigma_i X_i)
\]

\[
\sum_{i=1}^{n} m_i y_i''(t) = \sum_{i=1}^{n} V_i(t) \quad (\Sigma_i Y_i)
\]

Denote by

\[M = \sum_{i=1}^{n} m_i\]

the total mass of the system, by

\[X(t) = \frac{1}{M} \sum_{i=1}^{n} m_i x_i(t) \quad \text{and} \quad Y(t) = \frac{1}{M} \sum_{i=1}^{n} m_i y_i(t)\]

the \(x\)- and \(y\)-coordinates of the centre of mass of the system at time \(t\) and by

\[H(t) = \sum_{i=1}^{n} H_i(t) \quad \text{and} \quad V(t) = \sum_{i=1}^{n} V_i(t)\]

the \(x\)- and \(y\)-coordinates of the total external force acting on the system at time \(t\). In this notation, the equations \((\Sigma_i X_i)\) and \((\Sigma_i Y_i)\) are

\[MX''(t) = H(t) \quad MY''(t) = V(t) \quad (2.3.4)\]

So the centre of mass of the system moves just like a single particle of mass \(M\) subject to the total external force.

Now multiply equation \((Y_i)\) by \(x_i(t)\), subtract from it equation \((X_i)\) multiplied by \(y_i(t)\), and sum over \(i\). This gives the equation \(\sum_i [x_i(t) (Y_i) - y_i(t) (X_i)]\):

\[
\sum_{i=1}^{n} m_i \left[ x_i(t) y_i''(t) - y_i(t) x_i''(t) \right] = \sum_{i=1}^{n} \left[ x_i(t) V_i(t) - y_i(t) H_i(t) \right] + \sum_{1 \leq i, j \leq n} \left[ x_i(t) V_{ij}(t) - y_i(t) H_{ij}(t) \right]
\]

---

230
By the assumption (A2)
\[ x_1(t)V_{1,2}(t) - y_1(t)H_{1,2}(t) = x_1(t)M_{1,2}(t)[y_1(t) - y_2(t)] - y_1(t)M_{1,2}(t)[x_1(t) - x_2(t)] \]
\[ = M_{1,2}(t)[y_1(t)x_2(t) - x_1(t)y_2(t)] \]
\[ x_2(t)V_{2,1}(t) - y_2(t)H_{2,1}(t) = x_2(t)M_{2,1}(t)[y_2(t) - y_1(t)] - y_2(t)M_{2,1}(t)[x_2(t) - x_1(t)] \]
\[ = M_{2,1}(t)[-y_1(t)x_2(t) + x_1(t)y_2(t)] \]
\[ = M_{1,2}(t)[-y_1(t)x_2(t) + x_1(t)y_2(t)] \]

So the \( i = 1, j = 2 \) term in \( \sum_{1 \leq i, j \leq n} [x_i(t)V_{ij}(t) - y_i(t)H_{ij}(t)] \) exactly cancels the \( i = 2, j = 1 \) term. In this way all of the terms in \( \sum_{1 \leq i, j \leq n} [x_i(t)V_{ij}(t) - y_i(t)H_{ij}(t)] \) with \( i \neq j \) cancel. Each term with \( i = j \) is exactly zero. So \( \sum_{1 \leq i, j \leq n} [x_i(t)V_{ij}(t) - y_i(t)H_{ij}(t)] = 0 \) and
\[ \sum_{i=1}^{n} m_i [x_i(t)y''(t) - y_i(t)x''(t)] = \sum_{i=1}^{n} [x_i(t)V_i(t) - y_i(t)H_i(t)] \]

Define
\[ L(t) = \sum_{i=1}^{n} m_i [x_i(t)y'(t) - y_i(t)x'(t)] \]
\[ T(t) = \sum_{i=1}^{n} [x_i(t)V_i(t) - y_i(t)H_i(t)] \]

In this notation
\[ \frac{d}{dt}L(t) = T(t) \quad (2.3.5) \]

- Equation \((2.3.5)\) plays the role of Newton’s law of motion for rotational motion.
- \(T(t)\) is called the torque and plays the role of “rotational force”.
- \(L(t)\) is called the angular momentum (about the origin) and is a measure of the rate at which the piece of wood is rotating.

For example, if a particle of mass \( m \) is traveling in a circle of radius \( r \), centred on the origin, at \( \omega \) radians per unit time, then \( x(t) = r \cos(\omega t), y(t) = r \sin(\omega t) \) and
\[ m[x(t)y'(t) - y(t)x'(t)] = m[r \cos(\omega t) r \omega \cos(\omega t) - r \sin(\omega t) (-r \omega \sin(\omega t))] \]
\[ = mr^2 \omega \]
is proportional to \( \omega \), which is the rate of rotation about the origin.
In any event, in order for the piece of wood to remain stationary, that is to have \( x_i(t) \) and \( y_i(t) \) be constant for all \( 1 \leq i \leq n \), we need to have

\[
X''(y) = Y''(t) = L(t) = 0
\]

and then equations (2.3.4) and (2.3.5) force

\[
H(t) = V(t) = T(t) = 0
\]

Now suppose that the piece of wood is a seesaw that is long and thin and is lying on the \( x \)-axis, supported on a fulcrum at \( x = p \). Then every \( y_i = 0 \) and the torque simplifies to

\[
T(t) = \sum_{i=1}^{n} x_i(t)V_i(t).
\]

The forces consist of

- gravity, \( m_i g \), acting downwards on particle number \( i \), for each \( 1 \leq i \leq n \) and the
- force \( F \) imposed by the fulcrum that is pushing straight up on the particle at \( x = p \).

So

- The net vertical force is

\[
V(t) = F - \sum_{i=1}^{n} m_i g = F - Mg.
\]

If the seesaw is to remain stationary, this must be zero so that \( F = Mg \).

- The total torque (about the origin) is

\[
T = Fp - \sum_{i=1}^{n} m_i g x_i = Mgp - \sum_{i=1}^{n} m_i g x_i.
\]

If the seesaw is to remain stationary, this must also be zero and the fulcrum must be placed at

\[
.p = \frac{1}{M} \sum_{i=1}^{n} m_i x_i \quad (2.3.6)
\]

which is the centre of mass of the piece of wood.

2.4 Separable Differential Equations

A differential equation is an equation for an unknown function that involves the derivative of the unknown function. Differential equations play a central role in modelling a huge number of different phenomena. Here is a table giving a bunch of named differential equations and what they are used for. It is far from complete.
Newton’s Law of Motion describes motion of particles
Maxwell’s equations describes electromagnetic radiation
Navier–Stokes equations describes fluid motion
Heat equation describes heat flow
Wave equation describes wave motion
Schrödinger equation describes atoms, molecules and crystals
Stress-strain equations describes elastic materials
Black–Scholes models used for pricing financial options
Predator–prey equations describes ecosystem populations
Einstein’s equations connects gravity and geometry
Ludwig–Jones–Holling’s equation models spruce budworm/Balsam fir ecosystem
Zeeman’s model models heart beats and nerve impulses
Sherman–Rinzel–Keizer model for electrical activity in Pancreatic β–cells
Hodgkin–Huxley equations models nerve action potentials

We are just going to scratch the surface of the study of differential equations. Most universities offer half a dozen different undergraduate courses on various aspects of differential equations. We will just look at one special, but important, type of equation.

### 2.4.1 **Separate and integrate**

A separable differential equation is an equation for a function $y(x)$ of the form

$$\frac{dy}{dx}(x) = f(x) g(y(x))$$

We’ll start by developing a recipe for solving separable differential equations. Then we’ll look at many examples. Usually one suppresses the argument of $y(x)$ and writes the equation

$$\frac{dy}{dx} = f(x) g(y)$$

and solves such an equation by cross multiplying/dividing to get all of the $y$’s, including the $dy$ on one side of the equation and all of the $x$’s, including the $dx$, on the other side of the equation.

$$\frac{dy}{g(y)} = f(x) \, dx$$

19 Look at the right hand side of the equation. The $x$–dependence is separated from the $y$–dependence. That’s the reason for the name “separable”.

---

233
APPLICATIONS OF INTEGRATION

2.4 SEPARABLE DIFFERENTIAL EQUATIONS

We are of course assuming that \( g(y) \) is nonzero.) Then you integrate both sides

\[
\int \frac{dy}{g(y)} = \int f(x) \, dx \tag{2.4.1}
\]

This looks illegal, and indeed is illegal — \( \frac{dy}{dx} \) is not a fraction. But we’ll now see that the answer is still correct. This procedure is simply a mnemonic device to help you remember that answer. Let \( G(y) \) be an antiderivative of \( \frac{1}{g(y)} \) (i.e. \( G'(y) = \frac{1}{g(y)} \)) and \( F(x) \) be an antiderivative of \( f(x) \) (i.e. \( F'(x) = f(x) \)). If we reinstate the argument of \( y \), (2.4.1) is

\[
G(y(x)) = F(x) + C \tag{2.4.2}
\]

To check that a function \( y(x) \) obeys \( \frac{dy}{dx}(x) = f(x) \, g(y(x)) \) if and only if it obeys (2.4.2), just differentiate both sides of (2.4.2) with respect to \( x \). By the chain rule

\[
G(y(x)) = F(x) + C \iff G'(y(x)) \, y'(x) = F'(x) \iff \frac{y'(x)}{g(y(x))} = f(x)
\]

\[
\iff y'(x) = f(x) \, g(y(x))
\]

(We have again assumed that \( g(y) \) is nonzero.)

Observe that the solution (2.4.2) contains an arbitrary constant, \( C \). The value of this arbitrary constant cannot be determined by the differential equation. You need additional data to determine it. Often this data consists of the value of the unknown function for one value of \( x \). That is, often the problem you have to solve is of the form

\[
\frac{dy}{dx}(x) = f(x) \, g(y(x)) \quad y(x_0) = y_0
\]

where \( f(x) \) and \( g(y) \) are given functions and \( x_0 \) and \( y_0 \) are given numbers. This type of problem is called an “initial value problem”. It is solved by first using the method above to find the general solution to the differential equation, including the arbitrary constant \( C \), and then using the “initial condition” \( y(x_0) = y_0 \) to determine the value of \( C \). We’ll see examples of this shortly.

**Example 2.4.2**

Let \( a \) and \( b \) be any two constants. We’ll now solve the family of differential equations

\[
\frac{dy}{dx} = a(y - b)
\]

using our mnemonic device.

\[
\frac{dy}{y - b} = a \, dx \implies \int \frac{dy}{y - b} = \int a \, dx \implies \log |y - b| = ax + c \implies |y - b| = e^{ax+c} = e^c e^{ax}
\]

\[
y - b = Ce^{ax}
\]

where \( C \) is either \( +e^c \) or \( -e^c \). Note that as \( c \) runs over all real numbers, \( +e^c \) runs over all strictly positive real numbers and \( -e^c \) runs over all strictly negative real numbers. So, so far, \( C \) can be any real number except 0. But we were a bit sloppy here. We implicitly
assumed that \( y - b \) was nonzero, so that we could divide it across. None–the–less, the constant function \( y = b \), which corresponds to \( C = 0 \), is a perfectly good solution — when \( y \) is the constant function \( y = b \), both \( \frac{dy}{dx} \) and \( a(y - b) \) are zero. So the general solution to

\[
\frac{dy}{dx} = a(y - b)
\]

is

\[
y(x) = (y(0) - b)e^{ax} + b
\]

This is worth stating as a theorem.

**Theorem 2.4.3.**

Let \( a \) and \( b \) be constants. The differentiable function \( y(x) \) obeys the differential equation

\[
\frac{dy}{dx} = a(y - b)
\]

if and only if

\[
y(x) = (y(0) - b)e^{ax} + b
\]

**Example 2.4.4**

Solve \( \frac{dy}{dx} = y^2 \)

**Solution.** When \( y \neq 0 \),

\[
\frac{dy}{dx} = y^2 \implies \frac{dy}{y^2} = dx \implies \frac{y^{-1}}{-1} = x + C \implies y = -\frac{1}{x + C}
\]

When \( y = 0 \), this computation breaks down because \( \frac{dy}{y^2} \) contains a division by 0. We can check if the function \( y(x) = 0 \) satisfies the differential equation by just subbing it in:

\[
y(x) = 0 \implies y'(x) = 0, \ y(x)^2 = 0 \implies y'(x) = y(x)^2
\]

So \( y(x) = 0 \) is a solution and the full solution is

\[
y(x) = 0 \text{ or } y(x) = -\frac{1}{x + C}, \text{ for any constant } C
\]

**Example 2.4.5**

When a raindrop falls it increases in size so that its mass \( m(t) \), is a function of time \( t \). The
rate of growth of mass, i.e. \( \frac{dm}{dt} \), is \( km(t) \) for some positive constant \( k \). According to Newton’s law of motion, \( \frac{d}{dt}(mv) = gm \), where \( v \) is the velocity of the raindrop (with \( v \) being positive for downward motion) and \( g \) is the acceleration due to gravity. Find the terminal velocity, \( \lim_{t \to \infty} v(t) \), of a raindrop.

Solution. In this problem we have two unknown functions, \( m(t) \) and \( v(t) \), and two differential equations, \( \frac{dm}{dt} = km \) and \( \frac{d}{dt}(mv) = gm \). The first differential equation, \( \frac{dm}{dt} = km \), involves only \( m(t) \), not \( v(t) \), so we use it to determine \( m(t) \). By Theorem 2.4.3, with \( b = 0 \), \( a = k \), \( y \) replaced by \( m \) and \( x \) replaced by \( t \),

\[
\frac{dm}{dt} = km \implies m(t) = m(0)e^{kt}
\]

Now that we know \( m(t) \) (except for the value of the constant \( m(0) \)), we can substitute it into the second differential equation, which we can then use to determine the remaining unknown function \( v(t) \). Observe that the second equation, \( \frac{d}{dt}(mv) = gm = gm(0)e^{kt} \), tells that the derivative of the function \( y(t) = m(t)v(t) \) is \( gm(0)e^{kt} \). So \( y(t) \) is just an antiderivative of \( gm(0)e^{kt} \).

\[
\frac{dy}{dt} = gm = gm(0)e^{kt} \implies y(t) = \int gm(0)e^{kt} \, dt = gm(0)\frac{e^{kt}}{k} + C
\]

Now that we know \( y(t) = m(t)v(t) = m(0)e^{kt}v(t) \), we can get \( v(t) \) just by dividing out the \( m(0)e^{kt} \).

\[
y(t) = gm(0)\frac{e^{kt}}{k} + C \implies m(0)e^{kt}v(t) = gm(0)\frac{e^{kt}}{k} + C \implies v(t) = \frac{g}{k} + \frac{C}{m(0)e^{kt}}
\]

Our solution, \( v(t) \), contains two arbitrary constants, namely \( C \) and \( m(0) \). They will be determined by, for example, the mass and velocity at time \( t = 0 \). But since we are only interested in the terminal velocity \( \lim_{t \to \infty} v(t) \), we don’t need to know \( C \) and \( m(0) \). Since \( k > 0 \), \( \lim_{t \to \infty} \frac{C}{e^{kt}} = 0 \) and the terminal velocity \( \lim_{t \to \infty} v(t) = \frac{g}{k} \).

**Example 2.4.5**

A glucose solution is administered intravenously into the bloodstream at a constant rate \( r \). As the glucose is added, it is converted into other substances at a rate that is proportional to the concentration at that time. The concentration, \( C(t) \), of the glucose in the bloodstream at time \( t \) obeys the differential equation

\[
\frac{dC}{dt} = r - kC
\]

where \( k \) is a positive constant of proportionality.

(a) Express \( C(t) \) in terms of \( k \) and \( C(0) \).
Applications of Integration 2.4 Separable Differential Equations

(b) Find \( \lim_{t \to \infty} C(t) \).

Solution. (a) Since \( r - kC = -k\left(C - \frac{r}{k}\right) \), the given equation is

\[
\frac{dC}{dt} = -k\left(C - \frac{r}{k}\right)
\]

which is of the form solved in Theorem 2.4.3 with \( a = -k \) and \( b = \frac{r}{k} \). So the solution is

\[
C(t) = \frac{r}{k} + \left(C(0) - \frac{r}{k}\right)e^{-kt}
\]

(b) For any \( k > 0 \), \( \lim_{t \to \infty} e^{-kt} = 0 \). Consequently, for any \( C(0) \) and any \( k > 0 \), \( \lim_{t \to \infty} C(t) = \frac{r}{k} \).

We could have predicted this limit without solving for \( C(t) \). If we assume that \( C(t) \) approaches some equilibrium value \( C_e \) as \( t \) approaches infinity, then taking the limits of both sides of \( \frac{dC}{dt} = r - kC \) as \( t \to \infty \) gives

\[
0 = r - kC_e \implies C_e = \frac{r}{k}
\]

2.4.2 Optional — Carbon Dating

Scientists can determine the age of objects containing organic material by a method called carbon dating or radiocarbon dating\(^{20}\). The bombardment of the upper atmosphere by cosmic rays converts nitrogen to a radioactive isotope of carbon, \( ^{14}C \), with a half–life of about 5730 years. Vegetation absorbs carbon dioxide from the atmosphere through photosynthesis and animals acquire \( ^{14}C \) by eating plants. When a plant or animal dies, it stops replacing its carbon and the amount of \( ^{14}C \) begins to decrease through radioactive decay. Therefore the level of radioactivity also decreases. More precisely, let \( Q(t) \) denote the amount of \( ^{14}C \) in the plant or animal \( t \) years after it dies. The number of radioactive decays per unit time, at time \( t \), is proportional to the amount of \( ^{14}C \) present at time \( t \), which is \( Q(t) \). Thus

\[
\frac{dQ}{dt}(t) = -kQ(t)
\]

Here \( k \) is a constant of proportionality that is determined by the half–life. We shall explain what half–life is, and also determine the value of \( k \), in Example 2.4.7, below.

Before we do so, let’s think about the sign in (2.4.3).

- Recall that \( Q(t) \) denotes a quantity, namely the amount of \( ^{14}C \) present at time \( t \). There cannot be a negative amount of \( ^{14}C \). Nor can this quantity be zero. (We would not use carbon dating when there is no \( ^{14}C \) present.) Consequently, \( Q(t) > 0 \).

\(^{20}\) Willard Libby, of Chicago University was awarded the Nobel Prize in Chemistry in 1960, for developing radiocarbon dating.
• As the time \( t \) increases, \( Q(t) \) decreases, because \(^{14}\text{C}\) is being continuously converted into \(^{14}\text{N}\) by radioactive decay\(^{21}\). Thus \( \frac{dQ}{dt}(t) < 0 \).

• The signs \( Q(t) > 0 \) and \( \frac{dQ}{dt}(t) < 0 \) are consistent with (2.4.3) provided the constant of proportionality \( k > 0 \).

• In (2.4.3), we chose to call the constant of proportionality \( "-k" \). We did so in order to make \( k > 0 \). We could just as well have chosen to call the constant of proportionality \( "K" \). That is, we could have replaced (2.4.3) by \( \frac{dQ}{dt}(t) = KQ(t) \). The constant of proportionality \( K \) would have to be negative, (and \( K \) and \( k \) would be related by \( K = -k \)).

**Example 2.4.7**

In this example, we determine the value of the constant of proportionality \( k \) in (2.4.3) that corresponds to the half–life of \(^{14}\text{C}\), which is 5730 years.

• Imagine that some plant or animal contains a quantity \( Q_0 \) of \(^{14}\text{C}\) at its time of death. Let’s choose the zero point of time \( t = 0 \) to be the instant that the plant or animal died.

• Denote by \( Q(t) \) the amount of \(^{14}\text{C}\) in the plant or animal \( t \) years after it died. Then \( Q(t) \) must obey both (2.4.3) and \( Q(0) = Q_0 \).

• Theorem 2.4.3, with \( b = 0 \) and \( a = -k \), then tells us that \( Q(t) = Q_0e^{-kt} \) for all \( t \geq 0 \).

• By definition, the half–life of \(^{14}\text{C}\) is the length of time that it takes for half of the \(^{14}\text{C}\) to decay. That is, the half–life \( t_{1/2} \) is determined by

\[
Q(t_{1/2}) = \frac{1}{2}Q(0) = \frac{1}{2}Q_0
\]

\[
Q_0e^{-kt_{1/2}} = \frac{1}{2}Q_0
\]

\[
e^{-kt_{1/2}} = \frac{1}{2}
\]

Taking the logarithm of both sides gives

\[
-kt_{1/2} = \log \frac{1}{2} = -\log 2 \implies k = \frac{\log 2}{t_{1/2}}
\]

Recall that, in this text, we use \( \log x \) to indicate the natural logarithm. That is,

\[
\log x = \log_e x = \log x
\]

We are told that, for \(^{14}\text{C}\), the half–life \( t_{1/2} = 5730 \), so

\[
k = \frac{\log 2}{5730} = 0.000121 \quad \text{to 6 decimal places}
\]

\(^{21}\) The precise transition is \(^{14}\text{C} \rightarrow ^{14}\text{N} + e^- + \bar{\nu}_e \) where \( e^- \) is an electron and \( \bar{\nu}_e \) is an electron neutrino.
2.4 Separable Differential Equations

Example 2.4.7

From the work in the above example we have accumulated enough new facts to make a corollary to Theorem 2.4.3.

**Corollary 2.4.8.**

The function \( Q(t) \) satisfies the equation

\[
\frac{dQ}{dt} = -kQ(t)
\]

if and only if

\[ Q(t) = Q(0) e^{-kt} \]

The half–life is defined to be the time \( t_{1/2} \) which obeys

\[ Q(t_{1/2}) = \frac{1}{2} Q(0) \]

The half–life is related to the constant \( k \) by

\[ t_{1/2} = \frac{\log 2}{k} \]

Now here is a typical problem that is solved using Corollary 2.4.8.

Example 2.4.9

A particular piece of parchment contains about 64% as much \(^{14}\)C as plants do today. Estimate the age of the parchment.

**Solution.** Let \( Q(t) \) denote the amount of \(^{14}\)C in the parchment \( t \) years after it was first created. By (2.4.3) and Example 2.4.7

\[
\frac{dQ}{dt}(t) = -kQ(t) \quad \text{with} \quad k = \frac{\log 2}{5730} = 0.000121
\]

By Corollary 2.4.8

\[ Q(t) = Q(0) e^{-kt} \]

The time at which \( Q(t) \) reaches 0.64 \( Q(0) \) is determined by

\[
\begin{align*}
Q(t) &= 0.64 Q(0) \\
Q(0) e^{-kt} &= 0.64 Q(0) \\
e^{-kt} &= 0.64 \\
-k t &= \log 0.64 \\
t &= \frac{\log 0.64}{-k} = \frac{\log 0.64}{-0.000121} = 3700 \quad \text{to 2 significant digits}
\end{align*}
\]
Applications of Integration

2.4 Separable Differential Equations

That is, the parchment \(^{22}\) is about 37 centuries old.

We have stated that the half-life of \(^{14}\)C is 5730 years. How can this be determined? We can explain this using the following example.

Example 2.4.9

Example 2.4.10

A scientist in a B-grade science fiction film is studying a sample of the rare and fictitious element, implausium. With great effort he has produced a sample of pure implausium. The next day — 17 hours later — he comes back to his lab and discovers that his sample is now only 37% pure. What is the half-life of the element?

Solution. We can again set up our problem using Corollary 2.4.8. Let \(Q(t)\) denote the quantity of implausium at time \(t\), measured in hours. Then we know

\[
Q(t) = Q(0) \cdot e^{-kt}
\]

We also know that

\[
Q(17) = 0.37Q(0).
\]

That enables us to determine \(k\) via

\[
0.37 = e^{-17k}
\]

and so

\[
k = -\frac{\log 0.37}{17} = 0.05849
\]

We can then convert this to the half-life using Corollary 2.4.8:

\[
t_{1/2} = \frac{\log 2}{k} \approx 11.85 \text{ hours}
\]

While this example is entirely fictitious, one really can use this approach to measure the half-life of materials.

2.4.3 Optional — Newton’s Law of Cooling

Newton’s law of cooling says:

The rate of change of temperature of an object is proportional to the difference in temperature between the object and its surroundings. The temperature of the surroundings is sometimes called the ambient temperature.

\(^{22}\) The British Museum has an Egyptian mathematical text from the seventeenth century B.C.
If we denote by $T(t)$ the temperature of the object at time $t$ and by $A$ the temperature of its surroundings, Newton’s law of cooling says that there is some constant of proportionality, $K$, such that

$$\frac{dT}{dt}(t) = K[T(t) - A] \quad (2.4.4)$$

This mathematical model of temperature change works well when studying a small object in a large, fixed temperature, environment. For example, a hot cup of coffee in a large room. Let’s start by thinking a little about the sign of the constant of proportionality. At any time $t$, there are three possibilities.

- If $T(t) > A$, that is, if the body is warmer than its surroundings, we would expect heat to flow from the body into its surroundings and so we would expect the body to cool off so that $\frac{dT}{dt}(t) < 0$. For this expectation to be consistent with (2.4.4), we need $K < 0$.
- If $T(t) < A$, that is the body is cooler than its surroundings, we would expect heat to flow from the surroundings into the body and so we would expect the body to warm up so that $\frac{dT}{dt}(t) > 0$. For this expectation to be consistent with (2.4.4), we again need $K < 0$.
- Finally if $T(t) = A$, that is the body and its environment have the same temperature, we would not expect any heat to flow between the two and so we would expect that $\frac{dT}{dt}(t) = 0$. This does not impose any condition on $K$.

In conclusion, we would expect $K < 0$. Of course, we could have chosen to call the constant of proportionality $-k$, rather than $K$. Then the differential equation would be

$$\frac{dT}{dt} = -k(T - A)$$

and we would expect $k > 0$.

**Example 2.4.11**

The temperature of a glass of iced tea is initially 5°C. After 5 minutes, the tea has heated to 10°C in a room where the air temperature is 30°C.

(a) Determine the temperature as a function of time.

(b) What is the temperature after 10 minutes?

(c) Determine when the tea will reach a temperature of 20°C.

**Solution.** (a)

- Denote by $T(t)$ the temperature of the tea $t$ minutes after it was removed from the fridge, and let $A = 30$ be the ambient temperature.

- By Newton’s law of cooling,

$$\frac{dT}{dt} = K(T - A) = K(T - 30)$$

for some, as yet unknown, constant of proportionality $K$.

23 It does not work so well when the object is of a similar size to its surroundings since the temperature of the surroundings will rise as the object cools. It also fails when there are phase transitions involved — for example, an ice-cube melting in a warm room does not obey Newton’s law of cooling.
• By Theorem 2.4.3 with $a = K$ and $b = 30$,
\[
T(t) = [T(0) - 30] e^{Kt} + 30 = 30 - 25e^{Kt}
\]
since the initial temperature $T(0) = 5$.

• This solution is not complete because it still contains an unknown constant, namely $K$. We have not yet used the given data that $T(5) = 10$. We can use it to determine $K$. At $t = 5$,
\[
T(5) = 30 - 25e^{5K} = 10 \implies e^{5K} = \frac{20}{25} \implies 5K = \log \frac{20}{25}
\]
\[
\implies K = \frac{1}{5} \log \frac{4}{5} = -0.044629
\]
to six decimal places.

(b) To find the temperature at 10 minutes we can just use the solution we have determined above.

\[
T(10) = 30 - 25e^{10K}
\]
\[
= 30 - 25e^{10 \cdot \frac{1}{5} \log \frac{4}{5}}
\]
\[
= 30 - 25e^{2 \log \frac{4}{5}} = 30 - 25e^{\log \frac{16}{25}}
\]
\[
= 30 - 16 = 14^\circ
\]

(c) The temperature is $20^\circ$ when
\[
30 - 25e^{Kt} = 20 \implies e^{Kt} = \frac{10}{25} \implies Kt = \log \frac{10}{25}
\]
\[
\implies t = \frac{1}{K} \log \frac{2}{5} = 20.5 \text{ min}
\]
to one decimal place.

Example 2.4.11

Example 2.4.12

A dead body is discovered at 3:45pm in a room where the temperature is $20^\circ$C. At that time the temperature of the body is $27^\circ$C. Two hours later, at 5:45pm, the temperature of the body is $25.3^\circ$C. What was the time of death? Note that the normal (adult human) body temperature is $37^\circ$C.

Solution. We will assume that the body’s temperature obeys Newton’s law of cooling.

• Denote by $T(t)$ the temperature of the body at time $t$, with $t = 0$ corresponding to 3:45pm. We wish to find the time of death — call it $t_d$.

• There is a lot of data in the statement of the problem. We are told

(1) the ambient temperature: $A = 20$
(2) the temperature of the body when discovered: \( T(0) = 27 \)
(3) the temperature of the body 2 hours later: \( T(2) = 25.3 \)
(4) assuming the person was a healthy adult right up until he died, the temperature at the time of death: \( T(t_d) = 37 \).

- Theorem 2.4.3 with \( a = K \) and \( b = A = 20 \):
  \[
  T(t) = \left[ T(0) - A \right] e^{Kt} + A = 20 + 7e^{Kt}
  \]
  Two unknowns remain, \( K \) and \( t_d \).

- We can find the first, \( K \), by using the condition (3), which says \( T(2) = 25.3 \).
  \[
  25.3 = T(2) = 20 + 7e^{2K} \implies 7e^{2K} = 5.3 \implies 2K = \log \left( \frac{5.3}{7} \right) \implies K = \frac{1}{2} \log \left( \frac{5.3}{7} \right) = -0.139
  \]

- Finally, \( t_d \) is determined by the condition (4).
  \[
  37 = T(t_d) = 20 + 7e^{-0.139t_d} \implies e^{-0.139t_d} = \frac{17}{7} \implies -0.139t_d = \log \left( \frac{17}{7} \right)
  \implies t_d = -\frac{1}{0.139} \log \left( \frac{17}{7} \right) = -6.38
  \]
  to two decimal places. Now 6.38 hours is 6 hours and \( 0.38 \times 60 = 23 \) minutes. So the time of death was 6 hours and 23 minutes before 3:45pm, which is 9:22am.

Example 2.4.12

A slightly tricky example — we need to determine the ambient temperature from three measurements at different times.

Example 2.4.13

A glass of room-temperature water is carried out onto a balcony from an apartment where the temperature is 22°C. After one minute the water has temperature 26°C and after two minutes it has temperature 28°C. What is the outdoor temperature?

Solution. We will assume that the temperature of the thermometer obeys Newton’s law of cooling.

- Let \( A \) be the outdoor temperature and \( T(t) \) be the temperature of the water \( t \) minutes after it is taken outside.

- By Newton’s law of cooling,
  \[
  T(t) = A + (T(0) - A)e^{Kt}
  \]
  Theorem 2.4.3 with \( a = K \) and \( b = A \). Notice there are 3 unknowns here — \( A, T(0) \) and \( K \) — so we need three pieces of information to find them all.

- We are told \( T(0) = 22 \), so
  \[
  T(t) = A + (22 - A)e^{Kt}.
  \]
• We are also told \( T(1) = 26 \), which gives

\[
26 = A + (22 - A)e^K
\]

rearrange things

\[
e^K = \frac{26 - A}{22 - A}
\]

• Finally, \( T(2) = 28 \), so

\[
28 = A + (22 - A)e^{2K}
\]

rearrange

but \( e^K = \frac{26 - A}{22 - A} \), so

\[
(26 - A)^2 = (28 - A)(22 - A)
\]

multiply through by \((22 - A)^2\)

We can expand out both sides and collect up terms to get

\[
(26^2) - 52A + A^2 = (28 	imes 22) - 50A + A^2
\]

\[
= 676 = 616
\]

\[
60 = 2A
\]

\[
30 = A
\]

So the temperature outside is \( 30^\circ \).

2.4.4 Optional — Population Growth

Suppose that we wish to predict the size \( P(t) \) of a population as a function of the time \( t \). In the most naive model of population growth, each couple produces \( \beta \) offspring (for some constant \( \beta \)) and then dies. Thus over the course of one generation \( \beta \frac{P(t)}{2} \) children are produced and \( P(t) \) parents die so that the size of the population grows from \( P(t) \) to

\[
P(t + t_g) = P(t) + \beta \frac{P(t)}{2} - P(t) = \frac{\beta}{2} P(t)
\]

where \( t_g \) denotes the lifespan of one generation. The rate of change of the size of the population per unit time is

\[
\frac{P(t + t_g) - P(t)}{t_g} = \frac{1}{t_g} \left[ \frac{\beta}{2} P(t) - P(t) \right] = bP(t)
\]

where \( b = \frac{\beta - 2}{2t_g} \) is the net birthrate per member of the population per unit time. If we approximate

\[
\frac{P(t + t_g) - P(t)}{t_g} \approx \frac{dP}{dt}(t)
\]
we get the differential equation

\[ \frac{dP}{dt} = bP(t) \]  

(2.4.5)

By Corollary 2.4.8, with \(-k\) replaced by \(b\),

\[ P(t) = P(0) \cdot e^{bt} \]  

(2.4.6)

This is called the Malthusian\(^{24}\) growth model. It is, of course, very simplistic. One of its main characteristics is that, since \(P(t + T) = P(0) \cdot e^{b(t+T)} = P(t) \cdot e^{bT}\), every time you add \(T\) to the time, the population size is multiplied by \(e^{bT}\). In particular, the population size doubles every \(\frac{\log 2}{b}\) units of time. The Malthusian growth model can be a reasonably good model only when the population size is very small compared to its environment\(^{25}\). A more sophisticated model of population growth, that takes into account the “carrying capacity of the environment” is considered below.

**Example 2.4.14**

In 1927 the population of the world was about 2 billion. In 1974 it was about 4 billion. Estimate when it reached 6 billion. What will the population of the world be in 2100, assuming the Malthusian growth model?

**Solution.** We follow our usual pattern for dealing with such problems.

- Let \(P(t)\) be the world’s population, in billions, \(t\) years after 1927. Note that 1974 corresponds to \(t = 1974 - 1927 = 47\).

- We are assuming that \(P(t)\) obeys equation (2.4.5). So, by (2.4.6)

\[ P(t) = P(0) \cdot e^{bt} \]

Notice that there are 2 unknowns here — \(b\) and \(P(0)\) — so we need two pieces of information to find them.

- We are told \(P(0) = 2\), so

\[ P(t) = 2 \cdot e^{bt} \]

- We are also told \(P(47) = 4\), which gives

\[ 4 = 2 \cdot e^{47b} \]

\[ e^{47b} = 2 \]

\[ b = \frac{\log 2}{47} = 0.0147 \]

to 3 decimal places

---

24 This is named after Rev. Thomas Robert Malthus. He described this model in a 1798 paper called “An essay on the principle of population”.

25 That is, the population has plenty of food and space to grow.
• We now know $P(t)$ completely, so we can easily determine the predicted population in 2100, i.e. at $t = 2100 - 1927 = 173$.

$$P(173) = 2e^{173b} = 2e^{173 \times 0.0147} = 12.7 \text{ billion}$$

• Finally, our crude model predicts that the population is 6 billion at the time $t$ that obeys

$$P(t) = 2e^{bt} = 6$$

$$e^{bt} = 3$$

clean up

take the log and clean up

$$t = \frac{\log 3}{b} = 47 \log 3 / \log 2 = 74.5$$

which corresponds to the middle of 2001.

Example 2.4.14

Logistic growth adds one more wrinkle to the simple population model. It assumes that the population only has access to limited resources. As the size of the population grows the amount of food available to each member decreases. This in turn causes the net birth rate $b$ to decrease. In the logistic growth model $b = b_0 \left(1 - \frac{P}{K}\right)$, where $K$ is called the carrying capacity of the environment, so that

$$P'(t) = b_0 \left(1 - \frac{P(t)}{K}\right) P(t)$$

This is a separable differential equation and we can solve it explicitly. We shall do so shortly. See Example 2.4.15 below. But, before doing that, we’ll see what we can learn about the behaviour of solutions to differential equations like this without finding formulae for the solutions. It turns out that we can learn a lot just by watching the sign of $P'(t)$. For concreteness, we’ll look at solutions of the differential equation

$$\frac{dP}{dt}(t) = \left(6000 - 3P(t)\right) P(t)$$

We’ll sketch the graphs of four functions $P(t)$ that obey this equation.

• For the first function, $P(0) = 0$.
• For the second function, $P(0) = 1000$.
• For the third function, $P(0) = 2000$.
• For the fourth function, $P(0) = 3000$.

The sketchs will be based on the observation that $(6000 - 3P) P = 3(2000 - P) P$

• is zero for $P = 0, 2000$,

---

26 The 2015 Revision of World Population, a publication of the United Nations, predicts that the world’s population in 2100 will be about 11 billion. But “about” covers a pretty large range. They give an 80% confidence interval running from 10 billion to 12.5 billion.

27 The world population really reached 6 billion in about 1999.
Applications of Integration

2.4 Separable Differential Equations

- is strictly positive for $0 < P < 2000$ and
- is strictly negative for $P > 2000$.

Consequently

$$\frac{dP}{dt}(t) = \begin{cases} 
0 & \text{if } P(t) = 0 \\
> 0 & \text{if } 0 < P(t) < 2000 \\
= 0 & \text{if } P(t) = 2000 \\
< 0 & \text{if } P(t) > 2000
\end{cases}$$

Thus if $P(t)$ is some function that obeys $\frac{dP}{dt}(t) = (6000 - 3P(t))P(t)$, then as the graph of $P(t)$ passes through the point $(t, P(t))$

the graph has

- slope zero, i.e. is horizontal, if $P(t) = 0$
- positive slope, i.e. is increasing, if $0 < P(t) < 2000$
- slope zero, i.e. is horizontal, if $P(t) = 2000$
- negative slope, i.e. is decreasing, if $0 < P(t) < 2000$

as illustrated in the figure

As a result,

- if $P(0) = 0$, the graph starts out horizontally. In other words, as $t$ starts to increase, $P(t)$ remains at zero, so the slope of the graph remains at zero. The population size remains zero for all time. As a check, observe that the function $P(t) = 0$ obeys $\frac{dP}{dt}(t) = (6000 - 3P(t))P(t)$ for all $t$.

- Similarly, if $P(0) = 2000$, the graph again starts out horizontally. So $P(t)$ remains at 2000 and the slope remains at zero. The population size remains 2000 for all time. Again, the function $P(t) = 2000$ obeys $\frac{dP}{dt}(t) = (6000 - 3P(t))P(t)$ for all $t$.

- If $P(0) = 1000$, the graph starts out with positive slope. So $P(t)$ increases with $t$. As $P(t)$ increases towards 2000, the slope $(6000 - 3P(t))P(t)$, while remaining positive, gets closer and closer to zero. As the graph approaches height 2000, it becomes more and more horizontal. The graph cannot actually cross from below 2000 to above 2000, because to do so it would have to have strictly positive slope for some value of $P$ above 2000, which is not allowed.

247
Applications of Integration

2.4 Separable Differential Equations

- If \( P(0) = 3000 \), the graph starts out with negative slope. So \( P(t) \) decreases with \( t \).
  As \( P(t) \) decreases towards 2000, the slope \((6000 - 3P(t))P(t)\), while remaining negative, gets closer and closer to zero. As the graph approaches height 2000, it becomes more and more horizontal. The graph cannot actually cross from above 2000 to below 2000, because to do so it would have to have negative slope for some value of \( P \) below 2000, which is not allowed.

These curves are sketched in the figure below. We conclude that for any initial population size \( P(0) \), except \( P(0) = 0 \), the population size approaches 2000 as \( t \to \infty \).

Now we’ll do an example in which we explicitly solve the logistic growth equation.

Example 2.4.15

In 1986, the population of the world was 5 billion and was increasing at a rate of 2\% per year. Using the logistic growth model with an assumed maximum population of 100 billion, predict the population of the world in the years 2000, 2100 and 2500.

Solution. Let \( y(t) \) be the population of the world, in billions of people, at time 1986 + \( t \).

The logistic growth model assumes

\[
y' = ay(K - y)
\]

where \( K \) is the carrying capacity and \( a = \frac{b_0}{K} \).

First we’ll determine the values of the constants \( a \) and \( K \) from the given data.

- We know that, if at time zero the population is below \( K \), then as time increases the population increases, approaching the limit \( K \) as \( t \) tends to infinity. So in this problem \( K \) is the maximum population. That is, \( K = 100 \).
- We are also told that, at time zero, the percentage rate of change of population, \( 100 \frac{y'}{y} \), is 2, so that, at time zero, \( \frac{y'}{y} = 0.02 \). But, from the differential equation, \( \frac{y'}{y} = a(K - y) \). Hence at time zero, \( 0.02 = a(100 - 5) \), so that \( a = \frac{2}{995} \).

We now know \( a \) and \( K \) and can solve the (separable) differential equation

\[
\frac{dy}{dt} = ay(K - y) \quad \Rightarrow \quad \frac{dy}{y(K - y)} = a dt \quad \Rightarrow \quad \int \frac{1}{K} \left[ \frac{1}{y} - \frac{1}{y - K} \right] dy = \int a dt
\]

\[
\Rightarrow \frac{1}{K} \left[ \log |y| - \log |y - K| \right] = at + C
\]

\[
\Rightarrow \log \frac{|y|}{|y - K|} = aKt + CK \quad \Rightarrow \quad \frac{|y|}{|y - K|} = De^{aKt}
\]
with $D = e^{CK}$. We know that $y$ remains between 0 and $K$, so that $\left| \frac{y}{K-y} \right| = \frac{y}{K-y}$ and our solution obeys

$$\frac{y}{K-y} = De^{aKt}$$

At this stage, we know the values of the constants $a$ and $K$, but not the value of the constant $D$. We are given that at $t = 0$, $y = 5$. Subbing in this, and the values of $K$ and $a$,

$$\frac{5}{100-5} = De^0 \implies D = \frac{5}{95}$$

So the solution obeys the algebraic equation

$$\frac{y}{100-y} = \frac{5}{95}e^{2t/95}$$

which we can solve to get $y$ as a function of $t$.

$$y = (100 - y) \frac{5}{95}e^{2t/95} \implies 95y = (500 - 5y)e^{2t/95}$$

$$\implies (95 + 5e^{2t/95})y = 500e^{2t/95}$$

$$\implies y = \frac{500e^{2t/95}}{95 + 5e^{2t/95}} = \frac{100e^{2t/95}}{19 + e^{2t/95}} = \frac{100}{1 + 19e^{-2t/95}}$$

Finally,

- In the year 2000, $t = 14$ and $y = \frac{100}{1 + 19e^{-28/95}} \approx 6.6$ billion.
- In the year 2100, $t = 114$ and $y = \frac{100}{1 + 19e^{-228/95}} \approx 36.7$ billion.
- In the year 2200, $t = 514$ and $y = \frac{100}{1 + 19e^{-1028/95}} \approx 100$ billion.

2.4.5 Optional — Mixing Problems

At time $t = 0$, where $t$ is measured in minutes, a tank with a 5-litre capacity contains 3 litres of water in which 1 kg of salt is dissolved. Fresh water enters the tank at a rate of 2 litres per minute and the fully mixed solution leaks out of the tank at the varying rate of $2t$ litres per minute.

(a) Determine the volume of solution $V(t)$ in the tank at time $t$.

(b) Determine the amount of salt $Q(t)$ in solution when the amount of water in the tank is at maximum.
Applications of Integration

2.4 Separable Differential Equations

Solution. (a) The rate of change of the volume in the tank, at time \( t \), is \( 2 - 2t \), because water is entering at a rate 2 and solution is leaking out at a rate \( 2t \). Thus

\[
\frac{dV}{dt} = 2 - 2t \implies dV = (2 - 2t) \, dt \implies V = \int (2 - 2t) \, dt = 2t - t^2 + C
\]

at least until \( V(t) \) reaches either the capacity of the tank or zero. When \( t = 0 \), \( V = 3 \) so \( C = 3 \) and \( V(t) = 3 + 2t - t^2 \). Observe that \( V(t) \) is at a maximum when \( \frac{dV}{dt} = 2 - 2t = 0 \), or \( t = 1 \).

(b) In the very short time interval from time \( t \) to time \( t + dt \), \( 2t \, dt \) litres of brine leaves the tank. That is, the fraction \( \frac{2t \, dt}{V(t)} \) of the total salt in the tank, namely \( Q(t) \), leaves. Thus salt is leaving the tank at the rate

\[
\frac{Q(t) \frac{2t \, dt}{V(t)}}{dt} = \frac{2t Q(t)}{3 + 2t - t^2} \text{ kilograms per minute}
\]

so

\[
\frac{dQ}{dt} = -\frac{2t Q(t)}{3 + 2t - t^2} \implies \frac{dQ}{Q} = -\frac{2t}{3 + 2t - t^2} \, dt = -\frac{2t}{(3 - t)(1 + t)} \, dt = \left[ \frac{3/2}{t - 3} + \frac{1/2}{t + 1} \right] \, dt
\]

\[
\implies \log Q = \frac{3}{2} \log |t - 3| + \frac{1}{2} \log |t + 1| + C
\]

We are interested in the time interval \( 0 \leq t \leq 1 \). In this time interval \( |t - 3| = 3 - t \) and \( |t + 1| = t + 1 \) so

\[
\log Q = \frac{3}{2} \log(3 - t) + \frac{1}{2} \log(t + 1) + C
\]

At \( t = 0 \), \( Q \) is 1 so

\[
\log 1 = \frac{3}{2} \log(3 - 0) + \frac{1}{2} \log(0 + 1) + C \implies C = \log 1 - \frac{3}{2} \log 3 - \frac{1}{2} \log 1 = -\frac{3}{2} \log 3
\]

At \( t = 1 \)

\[
\log Q = \frac{3}{2} \log(3 - 1) + \frac{1}{2} \log(1 + 1) - \frac{3}{2} \log 3 = 2 \log 2 - \frac{3}{2} \log 3 = \log 4 - \log 3^{3/2}
\]

so \( Q = \frac{4}{3^{3/2}} \).

---

Example 2.4.17

A tank contains 1500 liters of brine with a concentration of 0.3 kg of salt per liter. Another brine solution, this with a concentration of 0.1 kg of salt per liter is poured into the tank at a rate of 20 li/min. At the same time, 20 li/min of the solution in the tank, which is stirred continuously, is drained from the tank.

(a) How many kilograms of salt will remain in the tank after half an hour?
(b) How long will it take to reduce the concentration to 0.2 kg/li?

Solution. Denote by $Q(t)$ the amount of salt in the tank at time $t$. In a very short time interval $dt$, the incoming solution adds $20\, dt$ liters of a solution carrying 0.1 kg/li. So the incoming solution adds $0.1 \times 20\, dt = 2\, dt$ kg of salt. In the same time interval $20\, dt$ liters is drained from the tank. The concentration of the drained brine is $\frac{Q(t)}{1500}$. So $\frac{Q(t)}{1500} \times 20\, dt$ kg were removed. All together, the change in the salt content of the tank during the short time interval is

$$dQ = 2\, dt - \frac{Q(t)}{1500} \times 20\, dt = \left(2 - \frac{Q(t)}{75}\right) dt$$

The rate of change of salt content per unit time is

$$\frac{dQ}{dt} = 2 - \frac{Q(t)}{75} = -\frac{1}{75} (Q(t) - 150)$$

The solution of this equation is

$$Q(t) = \{Q(0) - 150\} e^{-t/75} + 150$$

by Theorem 2.4.3, with $a = -\frac{1}{75}$ and $b = 150$. At time 0, $Q(0) = 1500 \times 0.3 = 450$. So

$$Q(t) = 150 + 300e^{-t/75}$$

(a) At $t = 30$

$$Q(30) = 150 + 300e^{-30/75} = 351.1 \text{ kg}$$

(b) $Q(t) = 0.2 \times 1500 = 300 \text{ kg}$ is achieved when

$$150 + 300e^{-t/75} = 300 \implies 300e^{-t/75} = 150 \implies e^{-t/75} = 0.5$$

$$\implies -\frac{t}{75} = \log(0.5) \implies t = -75\log(0.5) = 51.99 \text{ min}$$

2.4.6 Optional — Interest on Investments

Suppose that you deposit $P$ in a bank account at time $t = 0$. The account pays $r\%$ interest per year compounded $n$ times per year.

- The first interest payment is made at time $t = \frac{1}{n}$. Because the balance in the account during the time interval $0 < t < \frac{1}{n}$ is $P$ and interest is being paid for (\frac{1}{n})^{th}$ of a year, that first interest payment is $\frac{1}{n} \times \frac{r}{100} \times P$. After the first interest payment, the balance in the account is $P + \frac{1}{n} \times \frac{r}{100} \times P = \left(1 + \frac{r}{100n}\right) P$.
- The second interest payment is made at time $t = \frac{2}{n}$. Because the balance in the account during the time interval $\frac{1}{n} < t < \frac{2}{n}$ is $(1 + \frac{r}{100n})P$ and interest is being paid for (\frac{1}{n})^{th}$ of a year, the second interest payment is $\frac{1}{n} \times \frac{r}{100} \times (1 + \frac{r}{100n})P$. After the second interest payment, the balance in the account is $(1 + \frac{r}{100n})P + \frac{1}{n} \times \frac{r}{100} \times (1 + \frac{r}{100n})P = (1 + \frac{r}{100n})^2 P$. 

---

251
Applications of Integration 2.4 Separable Differential Equations

- And so on.

In general, at time $t = \frac{m}{n}$ (just after the $m^{th}$ interest payment), the balance in the account is

$$B(t) = \left(1 + \frac{r}{100n}\right)^m P = \left(1 + \frac{r}{100n}\right)^{nt} P \quad (2.4.7)$$

Three common values of $n$ are 1 (interest is paid once a year), 12 (i.e. interest is paid once a month) and 365 (i.e. interest is paid daily). The limit $n \to \infty$ is called continuous compounding. Under continuous compounding, the balance at time $t$ is

$$B(t) = \lim_{n \to \infty} \left(1 + \frac{r}{100n}\right)^{nt} P$$

You may have already seen the limit

$$\lim_{x \to 0} (1 + x)^{a/x} = e^a \quad (2.4.8)$$

If so, you can evaluate $B(t)$ by applying (2.4.8) with $x = \frac{r}{100n}$ and $a = \frac{rt}{100}$ (so that $\frac{a}{x} = nt$). As $n \to \infty$, $x \to 0$ so that

$$B(t) = \lim_{n \to \infty} \left(1 + \frac{r}{100n}\right)^{nt} P = \lim_{x \to 0} (1 + x)^{a/x} P = e^a P = e^{rt/100} P \quad (2.4.9)$$

If you haven’t seen (2.4.8) before, that’s OK. In the following example, we rederive (2.4.9) using a differential equation instead of (2.4.8).

Example 2.4.18

Suppose, again, that you deposit $P$ in a bank account at time $t = 0$, and that the account pays $r\%$ interest per year compounded $n$ times per year, and denote by $B(t)$ the balance at time $t$. Suppose that you have just received an interest payment at time $t$. Then the next interest payment will be made at time $t + \frac{1}{n}$ and will be $\frac{1}{n} \times \frac{r}{100} \times B(t)$ = $\frac{r}{100n} B(t)$. So, calling $\frac{1}{n} = h$,

$$B(t + h) = B(t) + \frac{r}{100} B(t)h \quad \text{or} \quad \frac{B(t + h) - B(t)}{h} = \frac{r}{100} B(t)$$

To get continuous compounding we take the limit $n \to \infty$ or, equivalently, $h \to 0$. This gives

$$\lim_{h \to 0} \frac{B(t + h) - B(t)}{h} = \frac{r}{100} B(t) \quad \text{or} \quad \frac{dB}{dt}(t) = \frac{r}{100} B(t)$$

By Theorem 2.4.3, with $a = \frac{r}{100}$ and $b = 0$, (or Corollary 2.4.8 with $k = -\frac{r}{100}$),

$$B(t) = e^{rt/100} B(0) = e^{rt/100} P$$

once again.

Example 2.4.18

28 There are banks that advertise continuous compounding. You can find some by googling “interest is compounded continuously and paid”
Applications of Integration

2.4 Separable Differential Equations

Example 2.4.19

(a) A bank advertises that it compounds interest continuously and that it will double your money in ten years. What is the annual interest rate?

(b) A bank advertises that it compounds monthly and that it will double your money in ten years. What is the annual interest rate?

Solution. (a) Let the interest rate be \( r \)\% per year. If you start with \( P \), then after \( t \) years, you have \( Pe^{rt/100} \), under continuous compounding. This was (2.4.9). After 10 years you have \( Pe^{r/10} \). This is supposed to be \( 2P \), so

\[
P e^{r/10} = 2P \implies e^{r/10} = 2 \implies \frac{r}{10} = \log 2 \implies r = 10 \log 2 = 6.93\%
\]

(b) Let the interest rate be \( r \)\% per year. If you start with \( P \), then after \( t \) years, you have \( P(1 + \frac{r}{100 \times 12})^{12t} \), under monthly compounding. This was (2.4.7). After 10 years you have \( P(1 + \frac{r}{100 \times 12})^{120} \). This is supposed to be \( 2P \), so

\[
P(1 + \frac{r}{100 \times 12})^{120} = 2P \implies (1 + \frac{r}{1200})^{120} = 2 \implies 1 + \frac{r}{1200} = 2^{1/120}
\]

\[
\implies \frac{r}{1200} = 2^{1/120} - 1 \implies r = 1200(2^{1/120} - 1) = 6.95\%
\]

Example 2.4.20

A 25 year old graduate of UBC is given $50,000 which is invested at 5\% per year compounded continuously. The graduate also intends to deposit money continuously at the rate of $2000 per year.

(a) Find a differential equation that \( A(t) \) obeys, assuming that the interest rate remains 5\%.

(b) Determine the amount of money in the account when the graduate is 65.

(c) At age 65, the graduate will start withdrawing money continuously at the rate of \( W \) dollars per year. If the money must last until the person is 85, what is the largest possible value of \( W \)?

Solution. (a) Let’s consider what happens to \( A \) over a very short time interval from time \( t \) to time \( t + \Delta t \). At time \( t \) the account balance is \( A(t) \). During the (really short) specified time interval the balance remains very close to \( A(t) \) and so earns interest of \( \frac{5}{100} \times \Delta t \times A(t) \). During the same time interval, the graduate also deposits an additional $2000\Delta t$. So

\[
A(t + \Delta t) \approx A(t) + 0.05A(t)\Delta t + 2000\Delta t \implies \frac{A(t + \Delta t) - A(t)}{\Delta t} \approx 0.05A(t) + 2000
\]
In the limit $\Delta t \to 0$, the approximation becomes exact and we get

$$\frac{dA}{dt} = 0.05A + 2000$$

(b) The amount of money at time $t$ obeys

$$\frac{dA}{dt} = 0.05A(t) + 2000 = 0.05(A(t) + 40,000)$$

So by Theorem 2.4.3 (with $a = 0.05$ and $b = -40,000$),

$$A(t) = (A(0) + 40,000)e^{0.05t} - 40,000$$

At time 0 (when the graduate is 25), $A(0) = 50,000$, so the amount of money at time $t$ is

$$A(t) = 90,000e^{0.05t} - 40,000$$

In particular, when the graduate is 65 years old, $t = 40$ and

$$A(40) = 90,000e^{0.05\times40} - 40,000 = 625,015.05$$

(c) When the graduate stops depositing money and instead starts withdrawing money at a rate $W$, the equation for $A$ becomes

$$\frac{dA}{dt} = 0.05A - W = 0.05(A - 20W)$$

assuming that the interest rate remains 5%. This time, Theorem 2.4.3 (with $a = 0.05$ and $b = 20W$) gives

$$A(t) = (A(0) - 20W)e^{0.05t} + 20W$$

If we now reset our clock so that $t = 0$ when the graduate is 65, $A(0) = 625,015.05$. So the amount of money at time $t$ is

$$A(t) = 20W + e^{0.05t}(625,015.05 - 20W)$$

We want the account to be depleted when the graduate is 85. So, we want $A(20) = 0$. This is the case if

$$20W + e^{0.05\times20}(625,015.05 - 20W) = 0 \implies 20W + e^{625,015.05 - 20W} = 0$$

$$\implies 20(e - 1)W = 625,015.05e$$

$$\implies W = \frac{625,015.05e}{20(e - 1)} = 49,437.96$$

---

Example 2.4.20
You have probably\textsuperscript{1} learned about Taylor polynomials and, in particular, that

\[ e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + E_n(x) \]

where $E_n(x)$ is the error introduced when you approximate $e^x$ by its Taylor polynomial of degree $n$. You may\textsuperscript{2} have even seen a formula for $E_n(x)$. We are now going to ask what happens as $n$ goes to infinity? Does the error go zero, giving an exact formula for $e^x$? We shall later see that it does and that

\[ e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!} \]

At this point we haven’t defined, or developed any understanding of, this infinite sum. How do we compute the sum of an infinite number of terms? Indeed, when does a sum of an infinite number of terms even make sense? Clearly we need to build up foundations to deal with these ideas. Along the way we shall also see other functions for which the corresponding error obeys $\lim_{n \to \infty} E_n(x) = 0$ for some values of $x$ and not for other values of $x$.

To motivate the next section, consider using the above formula with $x = 1$ to compute the number $e$:

\[ e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!} \]

As we stated above, we don’t yet understand what to make of this infinite number of terms, but we might try to sneak up on it by thinking about what happens as we take

\textsuperscript{1} If you took Mathematics 100, 180, 104 or 184 then you definitely learned about Taylor polynomials. Now would be an excellent time to quickly read over your notes on the topic.

\textsuperscript{2} Again, if you took Mathematics 100, 180, 104 or 184 then you almost certainly did see a formula for the error $E_n(x)$ and even how to bound it.
more and more terms.

<table>
<thead>
<tr>
<th>Terms</th>
<th>Calculation</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 term</td>
<td>$1 = 1$</td>
<td></td>
</tr>
<tr>
<td>2 terms</td>
<td>$1 + 1 = 2$</td>
<td></td>
</tr>
<tr>
<td>3 terms</td>
<td>$1 + 1 + \frac{1}{2} = 2.5$</td>
<td></td>
</tr>
<tr>
<td>4 terms</td>
<td>$1 + 1 + \frac{1}{2} + \frac{1}{6} = 2.666666\ldots$</td>
<td></td>
</tr>
<tr>
<td>5 terms</td>
<td>$1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} = 2.708333\ldots$</td>
<td></td>
</tr>
<tr>
<td>6 terms</td>
<td>$1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120} = 2.716666\ldots$</td>
<td></td>
</tr>
</tbody>
</table>

By looking at the infinite sum in this way, we naturally obtain a sequence of numbers

\[ \{1, 2, 2.5, 2.666666, \ldots, 2.708333, \ldots, 2.716666, \ldots, \ldots\}. \]

The key to understanding the original infinite sum is to understand the behaviour of this sequence of numbers — in particularly, what do the numbers do as we go further and further? Does it settle down to a given limit?

### 3.1 Sequences

In the discussion above we used the term “sequence” without giving it a precise mathematical meaning. Let us rectify this now.

**Definition 3.1.1.**

A sequence is a list of infinitely many numbers with a specified order. It is denoted

\[ \{a_1, a_2, a_3, \ldots, a_n, \ldots\} \text{ or } \{a_n\} \text{ or } \{a_n\}_{n=1}^{\infty} \]

We will often specify a sequence by writing it more explicitly, like

\[ \left\{a_n = f(n)\right\}_{n=1}^{\infty} \]

where $f(n)$ is some function from the natural numbers to the real numbers.

---

3 You will notice a great deal of similarity between the results of the next section and “limits at infinity” which was covered last term.

4 For the more pedantic reader, here we mean a list of countably infinitely many numbers. The interested (pedantic or otherwise) reader should look up countable and uncountable sets.
Here are three sequences.

\[
\begin{align*}
\{1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots\} & \quad \text{or} \quad \left\{a_n = \frac{1}{n}\right\}_{n=1}^\infty \\
\{1, 2, 3, \ldots, n, \ldots\} & \quad \text{or} \quad \left\{a_n = n\right\}_{n=1}^\infty \\
\{1, -1, 1, -1, \ldots, (-1)^{n-1}, \ldots\} & \quad \text{or} \quad \left\{a_n = (-1)^{n-1}\right\}_{n=1}^\infty
\end{align*}
\]

It is not necessary that there be a simple explicit formula for the \(n\)th term of a sequence. For example the decimal digits of \(\pi\) is a perfectly good sequence

\[
\{3, 1, 4, 1, 5, 9, 2, 6, 5, 3, 5, 8, 9, 7, 9, 3, 2, 3, 8, 4, 6, 2, 6, 4, 3, 3, 8, \ldots\}
\]

but there is no simple formula\(^5\) for the \(n\)th digit.

Our primary concern with sequences will be the behaviour of \(a_n\) as \(n\) tends to infinity and, in particular, whether or not \(a_n\) “settles down” to some value as \(n\) tends to infinity.

**Definition 3.1.3.**

A sequence \(\{a_n\}_{n=1}^\infty\) is said to converge to the limit \(A\) if \(a_n\) approaches \(A\) as \(n\) tends to infinity. If so, we write

\[
\lim_{n \to \infty} a_n = A \quad \text{or} \quad a_n \to A \quad \text{as} \quad n \to \infty
\]

A sequence is said to converge if it converges to some limit. Otherwise it is said to diverge.

The reader should immediately recognise the similarity with limits at infinity

\[
\lim_{x \to \infty} f(x) = L \quad \text{if} \quad f(x) \to L \quad \text{as} \quad x \to \infty
\]

**Example 3.1.4**

Three of the four sequences in Example 3.1.2 diverge:

- The sequence \(\{a_n = n\}_{n=1}^\infty\) diverges because \(a_n\) grows without bound, rather than approaching some finite value, as \(n\) tends to infinity.

---

\(^5\) There is, however, a remarkable result due to Bailey, Borwein and Plouffe that can be used to compute the \(n\)th binary digit of \(\pi\) (i.e. writing \(\pi\) in base 2 rather than base 10) without having to work out the preceding digits.
• The sequence \( \{a_n = (-1)^{n-1}\}_{n=1}^{\infty} \) diverges because \( a_n \) oscillates between \(+1\) and \(-1\) rather than approaching a single value as \( n \) tends to infinity.

• The sequence of the decimal digits of \( \pi \) also diverges, though the proof that this is the case is a bit beyond us right now\(^6\).

The other sequence in Example 3.1.2 has \( a_n = \frac{1}{n} \). As \( n \) tends to infinity, \( \frac{1}{n} \) tends to zero. So

\[
\lim_{n \to \infty} \frac{1}{n} = 0
\]

Example 3.1.4

Example 3.1.5 \( \left( \lim_{n \to \infty} \frac{n}{2n+1} \right) \)

Here is a little less trivial example. To study the behaviour of \( \frac{n}{2n+1} \) as \( n \to \infty \), it is a good idea to write it as

\[
\frac{n}{2n+1} = \frac{1}{2 + \frac{1}{n}}
\]

As \( n \to \infty \), the \( \frac{1}{n} \) in the denominator tends to zero, so that the denominator \( 2 + \frac{1}{n} \) tends to 2 and \( \frac{1}{2 + \frac{1}{n}} \) tends to \( \frac{1}{2} \). So

\[
\lim_{n \to \infty} \frac{n}{2n+1} = \lim_{n \to \infty} \frac{1}{2 + \frac{1}{n}} = \frac{1}{2}
\]

Example 3.1.5

Notice that in this last example, we are really using techniques that we used before to study infinite limits like \( \lim_{x \to \infty} f(x) \). This experience can be easily transferred to dealing with \( \lim_{n \to \infty} \) limits by using the following result.

**Theorem 3.1.6.**

If

\[
\lim_{x \to \infty} f(x) = L
\]

and if \( a_n = f(n) \) for all positive integers \( n \), then

\[
\lim_{n \to \infty} a_n = L
\]

\(^6\) If the digits of \( \pi \) were to converge, then \( \pi \) would have to be a rational number. The irrationality of \( \pi \) (that it cannot be written as a fraction) was first proved by Lambert in 1761. Niven’s 1947 proof is more accessible and we invite the interested reader to use their favourite search engine to find step–by–step guides to that proof.
Example 3.1.7 \( \left( \lim_{n \to \infty} e^{-n} \right) \)

Set \( f(x) = e^{-x} \). Then \( e^{-n} = f(n) \) and

\[
\text{since } \lim_{x \to \infty} e^{-x} = 0 \quad \text{we know that} \quad \lim_{n \to \infty} e^{-n} = 0
\]

The bulk of the rules for the arithmetic of limits of functions that you already know also apply to the limits of sequences. That is, the rules you learned to work with limits such as \( \lim_{x \to \infty} f(x) \) also apply to limits like \( \lim_{n \to \infty} a_n \).

**Theorem 3.1.8 (Arithmetic of limits).**

Let \( A, B \) and \( C \) be real numbers and let the two sequences \( \{a_n\}_{n=1}^{\infty} \) and \( \{b_n\}_{n=1}^{\infty} \) converge to \( A \) and \( B \) respectively. That is, assume that

\[
\lim_{n \to \infty} a_n = A \quad \lim_{n \to \infty} b_n = B
\]

Then the following limits hold.

(a) \( \lim_{n \to \infty} [a_n + b_n] = A + B \)
   (The limit of the sum is the sum of the limits.)

(b) \( \lim_{n \to \infty} [a_n - b_n] = A - B \)
   (The limit of the difference is the difference of the limits.)

(c) \( \lim_{n \to \infty} Ca_n = CA. \)

(d) \( \lim_{n \to \infty} a_n b_n = AB \)
   (The limit of the product is the product of the limits.)

(e) If \( B \neq 0 \) then \( \lim_{n \to \infty} \frac{a_n}{b_n} = \frac{A}{B} \)
   (The limit of the quotient is the quotient of the limits provided the limit of the denominator is not zero.)

We use these rules to evaluate limits of more complicated sequences in terms of the limits of simpler sequences — just as we did for limits of functions.
Combining Examples 3.1.5 and 3.1.7,

\[
\lim_{n \to \infty} \left( \frac{n}{2n+1} + 7e^{-n} \right) = \lim_{n \to \infty} \frac{n}{2n+1} + \lim_{n \to \infty} 7e^{-n} \quad \text{by Theorem 3.1.8.a}
\]

\[
= \lim_{n \to \infty} \frac{n}{2n+1} + 7 \lim_{n \to \infty} e^{-n} \quad \text{by Theorem 3.1.8.c}
\]

\[
= \frac{1}{2} + 7 \cdot 0 \quad \text{by Examples 3.1.5 and 3.1.7}
\]

\[
= \frac{1}{2}
\]

There is also a squeeze theorem for sequences.

**Theorem 3.1.10 (Squeeze theorem).**

If \( a_n \leq c_n \leq b_n \) for all natural numbers \( n \), and if

\[
\lim_{n \to \infty} a_n = \lim_{n \to \infty} b_n = L
\]

then

\[
\lim_{n \to \infty} c_n = L
\]

In this example we use the squeeze theorem to evaluate

\[
\lim_{n \to \infty} \left[ 1 + \frac{\pi_n}{n} \right]
\]

where \( \pi_n \) is the \( n \)th decimal digit of \( \pi \). That is,

\[
\pi_1 = 3 \quad \pi_2 = 1 \quad \pi_3 = 4 \quad \pi_4 = 1 \quad \pi_5 = 5 \quad \pi_6 = 9 \quad \ldots
\]

We do not have a simple formula for \( \pi_n \). But we do know that

\[
0 \leq \pi_n \leq 9 \implies 0 \leq \frac{\pi_n}{n} \leq \frac{9}{n} \implies 1 \leq 1 + \frac{\pi_n}{n} \leq 1 + \frac{9}{n}
\]

and we also know that

\[
\lim_{n \to \infty} 1 = 1 \quad \lim_{n \to \infty} \left[ 1 + \frac{9}{n} \right] = 1
\]

So the squeeze theorem with \( a_n = 1 \), \( b_n = 1 + \frac{\pi_n}{n} \), and \( c_n = 1 + \frac{9}{n} \) gives

\[
\lim_{n \to \infty} \left[ 1 + \frac{\pi_n}{n} \right] = 1
\]
Finally, recall that we can compute the limit of the composition of two functions using continuity. In the same way, we have the following result:

**Theorem 3.1.12 (Continuous functions of limits).**

If \( \lim_{n \to \infty} a_n = L \) and if the function \( g(x) \) is continuous at \( L \), then

\[
\lim_{n \to \infty} g(a_n) = g(L)
\]

Example 3.1.13

Write \( \sin \frac{\pi n}{2n+1} = g\left( \frac{n}{2n+1} \right) \) with \( g(x) = \sin(\pi x) \). We saw, in Example 3.1.5 that

\[
\lim_{n \to \infty} \frac{n}{2n+1} = \frac{1}{2}
\]

Since \( g(x) = \sin(\pi x) \) is continuous at \( x = \frac{1}{2} \), which is the limit of \( \frac{n}{2n+1} \), we have

\[
\lim_{n \to \infty} \sin \frac{\pi n}{2n+1} = \lim_{n \to \infty} g\left( \frac{n}{2n+1} \right) = g\left( \frac{1}{2} \right) = \sin \frac{\pi}{2} = 1
\]

With this introduction to sequences and some tools to determine their limits, we can now return to the problem of understanding infinite sums.

### 3.2 Series

A series is a sum

\[ a_1 + a_2 + a_3 + \cdots + a_n + \cdots \]

of infinitely many terms. In summation notation, it is written

\[
\sum_{n=1}^{\infty} a_n
\]

You already have a lot of experience with series, though you might not realise it. When you write a number using its decimal expansion you are really expressing it as a series. Perhaps the simplest example of this is the decimal expansion of \( \frac{1}{3} \):  

\[ \frac{1}{3} = 0.3333\cdots \]
Recall that the expansion written in this way actually means

\[
\frac{3}{10} + \frac{3}{100} + \frac{3}{1000} + \frac{3}{10000} + \cdots = \sum_{n=1}^{\infty} \frac{3}{10^n}
\]

The summation index \( n \) is of course a dummy index. You can use any symbol you like (within reason) for the summation index.

\[
\sum_{n=1}^{\infty} \frac{3}{10^n} = \sum_{i=1}^{\infty} \frac{3}{10^i} = \sum_{j=1}^{\infty} \frac{3}{10^j} = \sum_{\ell=1}^{\infty} \frac{3}{10^\ell}
\]

A series can be expressed using summation notation in many different ways. For example the following expressions all represent the same series:

\[
\sum_{n=1}^{\infty} \frac{3}{10^n} = \sum_{j=1}^{\infty} \frac{3}{10^j} = \sum_{\ell=1}^{\infty} \frac{3}{10^\ell}
\]

We can get from the first line to the second line by substituting \( n = j - 1 \) — don’t forget to also change the limits of summation (so that \( n = 1 \) becomes \( j - 1 = 1 \) which is rewritten as \( j = 2 \)). To get from the first line to the third line, substitute \( n = \ell + 1 \) everywhere, including in the limits of summation (so that \( n = 1 \) becomes \( \ell + 1 = 1 \) which is rewritten as \( \ell = 0 \)).

Whenever you are in doubt as to what series a summation notation expression represents, it is a good habit to write out the first few terms, just as we did above.

Of course, at this point, it is not clear whether the sum of infinitely many terms adds up to a finite number or not. In order to make sense of this we will recast the problem in terms of the convergence of sequences (hence the discussion of the previous section). Before we proceed more formally let us illustrate the basic idea with a few simple examples.

**Example 3.2.1**

As we have just seen above the series \( \sum_{n=1}^{\infty} \frac{3}{10^n} \) is

\[
\sum_{n=1}^{\infty} \frac{3}{10^n} = \frac{3}{10} + \frac{3}{100} + \frac{3}{1000} + \cdots
\]
Notice that the \( n^{th} \) term in that sum is

\[
3 \times 10^{-n} = 0.\overline{00\ldots0}3
\]

So the sum of the first 5, 10, 15 and 20 terms in that series are

\[
\begin{align*}
\sum_{n=1}^{5} \frac{3}{10^n} &= 0.33333 \\
\sum_{n=1}^{10} \frac{3}{10^n} &= 0.3333333333 \\
\sum_{n=1}^{15} \frac{3}{10^n} &= 0.3333333333333333 \\
\sum_{n=1}^{20} \frac{3}{10^n} &= 0.33333333333333333333
\end{align*}
\]

It sure looks like that, as we add more and more terms, we get closer and closer to \( 0.\overline{3} = \frac{1}{3} \). So it is very reasonable\(^7\) to define \( \sum_{n=1}^{\infty} \frac{3}{10^n} \) to be \( \frac{1}{3} \).

Example 3.2.1

Every term in the series \( \sum_{n=1}^{\infty} 1 \) is exactly 1. So the sum of the first \( N \) terms is exactly \( N \). As we add more and more terms this grows unboundedly. So it is very reasonable to say that the series \( \sum_{n=1}^{\infty} 1 \) diverges.

The series

\[
\sum_{n=1}^{\infty} (-1)^n = \frac{n=1}{(-1)} + \frac{n=2}{1} + \frac{n=3}{(-1)} + \frac{n=4}{1} + \frac{n=5}{(-1)} + \ldots
\]

So the sum of the first \( N \) terms is 0 if \( N \) is even and \( -1 \) if \( N \) is odd. As we add more and more terms from the series, the sum alternates between 0 and \( -1 \) for ever and ever. So the sum of all infinitely many terms does not make any sense and it is again reasonable to say that the series \( \sum_{n=1}^{\infty} (-1)^n \) diverges.

Example 3.2.2

In the above examples we have tried to understand the series by examining the sum of the first few terms and then extrapolating as we add in more and more terms. That is, we tried to sneak up on the infinite sum by looking at the limit of (partial) sums of the first few terms. This approach can be made into a more formal rigorous definition. More precisely, to define what is meant by the infinite sum \( \sum_{n=1}^{\infty} a_n \), we approximate it by the sum of its first \( N \) terms and then take the limit as \( N \) tends to infinity.

\(^7\) Of course we are free to define the series to be whatever we want. The hard part is defining it to be something that makes sense and doesn’t lead to contradictions. We’ll get to a more systematic definition shortly.
3.2 Series

The \( N \)th partial sum of the series \( \sum_{n=1}^{\infty} a_n \) is the sum of its first \( N \) terms

\[
S_N = \sum_{n=1}^{N} a_n.
\]

The partial sums form a sequence \( \{S_N\}_{N=1}^{\infty} \). If this sequence of partial sums converges \( S_N \rightarrow S \) as \( N \rightarrow \infty \) then we say that the series \( \sum_{n=1}^{\infty} a_n \) converges to \( S \) and we write

\[
\sum_{n=1}^{\infty} a_n = S
\]

If the sequence of partial sums diverges, we say that the series diverges.

Definition 3.2.3.

Example 3.2.4 (Geometric Series)

Let \( a \) and \( r \) be any two fixed real numbers with \( a \neq 0 \). The series

\[
a + ar + ar^2 + \cdots + ar^n + \cdots = \sum_{n=0}^{\infty} ar^n
\]

is called the geometric series with first term \( a \) and ratio \( r \).

Notice that we have chosen to start the summation index at \( n = 0 \). That’s fine. The first term is the \( n = 0 \) term, which is \( ar^0 = a \). The second term is the \( n = 1 \) term, which is \( ar^1 = ar \). And so on. We could have also written the series \( \sum_{n=1}^{\infty} ar^{n-1} \). That’s exactly the same series — the first term is \( ar^{n-1}|_{n=1} = ar^{1-1} = a \), the second term is \( ar^{n-1}|_{n=2} = ar^{2-1} = ar \), and so on. Regardless of how we write the geometric series, \( a \) is the first term and \( r \) is the ratio between successive terms.

Geometric series have the extremely useful property that there is a very simple formula for their partial sums. Denote the partial sum by

\[
S_N = \sum_{n=0}^{N} ar^n = a + ar + ar^2 + \cdots + ar^N.
\]

---

8 It is actually quite common in computer science to think of 0 as the first integer. In that context, the set of natural numbers is defined to contain 0:

\[\mathbb{N} = \{0, 1, 2, \ldots\}\]

while the notation

\[\mathbb{Z}^+ = \{1, 2, 3, \ldots\}\]

is used to denote the (strictly) positive integers. Remember that in this text, as is more standard in mathematics, we define the set of natural numbers to be the set of (strictly) positive integers.

9 This reminds the authors of the paradox of Hilbert’s hotel. The hotel with an infinite number of rooms is completely full, but can always accommodate one more guest. The interested reader should use their favourite search engine to find more information on this.
The secret to evaluating this sum is to see what happens when we multiply it \( r \):
\[
    rS_N = r(a + ar + ar^2 + \cdots + ar^N)
    = ar + ar^2 + ar^3 + \cdots + ar^{N+1}
\]

Notice that this is almost the same\(^{10}\) as \( S_N \). The only differences are that the first term, \( a \), is missing and one additional term, \( ar^{N+1} \), has been tacked on the end. So
\[
    S_N = a + ar + ar^2 + \cdots + ar^N
    
    rS_N = ar + ar^2 + \cdots + ar^N + ar^{N+1}
\]

Hence taking the difference of these expressions cancels almost all the terms:
\[
(1 - r)S_N = a - ar^{N+1} = a(1 - r^{N+1})
\]

Provided \( r \neq 1 \) we can divide both side by \( 1 - r \) to isolate \( S_N \):
\[
    S_N = a \cdot \frac{1 - r^{N+1}}{1 - r}
\]

On the other hand, if \( r = 1 \), then
\[
    S_N = \underbrace{a + a + \cdots + a}_{N+1 \text{ terms}} = a(N + 1)
\]

So in summary:
\[
    S_N = \begin{cases} 
    a \frac{1 - r^{N+1}}{1 - r} & \text{if } r \neq 1 \\
    a(N + 1) & \text{if } r = 1 
    \end{cases}
\]

Now that we have this expression we can determine whether or not the series converges. If \(|r| < 1\), then \( r^{N+1} \) tends to zero as \( N \to \infty \), so that \( S_N \) converges to \( \frac{1}{1-r} \) as \( N \to \infty \) and
\[
    \sum_{n=0}^{\infty} ar^n = \frac{a}{1 - r} \quad \text{provided } |r| < 1.
\]

On the other hand if \(|r| \geq 1\), \( S_N \) diverges. To understand this divergence, consider the following 4 cases:

- If \( r > 1 \), then \( r^N \) grows to \( \infty \) as \( N \to \infty \).
- If \( r < -1 \), then the magnitude of \( r^N \) grows to \( \infty \), and the sign of \( r^N \) oscillates between + and −, as \( N \to \infty \).

\(^{10}\) One can find similar properties of other special series, that allow us, with some work, to cancel many terms in the partial sums. We will shortly see a good example of this. The interested reader should look up “creative telescoping” to see how this idea might be used more generally, though it is somewhat beyond this course.
• If \( r = +1 \), then \( N + 1 \) grows to \( \infty \) as \( N \to \infty \).

• If \( r = -1 \), then \( r^N \) just oscillates between \(+1\) and \(-1\) as \( N \to \infty \).

In each case the sequence of partial sums does not converge and so the series does not converge.

Example 3.2.4

Now that we know how to handle geometric series let’s return to Example 3.2.1.

Example 3.2.5 (Decimal Expansions)

The decimal expansion

\[
0.3333\ldots = \frac{3}{10} + \frac{3}{100} + \frac{3}{1000} + \frac{3}{10000} + \cdots = \sum_{n=1}^{\infty} \frac{3}{10^n}
\]

is a geometric series with the first term \( a = \frac{3}{10} \) and the ratio \( r = \frac{1}{10} \). So, by Example 3.2.4,

\[
0.3333\ldots = \sum_{n=1}^{\infty} \frac{3}{10^n} = \frac{3/10}{1 - 1/10} = \frac{3/10}{9/10} = \frac{1}{3}
\]

just as we would have expected.

We can push this idea further. Consider the repeating decimal expansion:

\[
0.16161616\ldots = \frac{16}{100} + \frac{16}{10000} + \frac{16}{1000000} + \cdots
\]

This is another geometric series with the first term \( a = \frac{16}{100} \) and the ratio \( r = \frac{1}{100} \). So, by Example 3.2.4,

\[
0.16161616\ldots = \sum_{n=1}^{\infty} \frac{16}{100^n} = \frac{16/100}{1 - 1/100} = \frac{16/100}{99/100} = \frac{1}{6}
\]

again, as expected. In this way any periodic decimal expansion converges to a ratio of two integers — that is, to a rational number.\(^\text{11}\)

Here is another more complicated example.

\[
0.12343434\ldots = \frac{12}{100} + \frac{34}{10000} + \frac{34}{1000000} + \cdots
\]

\[
= \frac{12}{100} + \sum_{n=2}^{\infty} \frac{34}{100^n}
\]

\[
= \frac{12}{100} + \frac{34}{10000} \frac{1}{1 - 1/100} \quad \text{by Example 3.2.4 with } a = \frac{34}{100^2} \text{ and } r = \frac{1}{100}
\]

\[
= \frac{12}{100} + \frac{34}{10000} \frac{99}{99}
\]

\[
= \frac{12}{100} + \frac{34}{9900}
\]

\[
= \frac{1222}{9900}
\]

\(^\text{11}\) We have included a (more) formal proof of this fact in the optional §3.7 at the end of this chapter. Proving that a repeating decimal expansion gives a rational number isn’t too hard. Proving the converse — that every rational number has a repeating decimal expansion is a little trickier, but we also do that in the same optional section.
Typically, it is quite difficult to write down a neat closed form expression for the partial sums of a series. Geometric series are very notable exceptions to this. Another family of series for which we can write down partial sums is called “telescoping series”. These series have the desirable property that many of the terms in the sum cancel each other out rendering the partial sums quite simple.

Example 3.2.6 (Telescoping Series)

In this example, we are going to study the series $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$. This is a rather artificial series that has been rigged to illustrate a phenomenon call “telescoping”. Notice that the $n^{\text{th}}$ term can be rewritten as

$$\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$$

and so we have

$$a_n = b_n - b_{n+1}$$

where $b_n = \frac{1}{n}$.

Because of this we get big cancellations when we add terms together. This allows us to get a simple formula for the partial sums of this series.

$$S_N = \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{N \cdot (N+1)}$$

$$= \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{N} - \frac{1}{N+1}\right)$$

The second term of each bracket exactly cancels the first term of the following bracket. So the sum “telescopes” leaving just

$$S_N = 1 - \frac{1}{N+1}$$

and we can now easily compute

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = \lim_{N \to \infty} S_N = \lim_{N \to \infty} \left(1 - \frac{1}{N+1}\right) = 1$$

Example 3.2.6

Well... this sort of series does show up when you start to look at the Maclaurin polynomial of functions like $(1 - x) \log(1 - x)$. So it is not totally artificial. At any rate, it illustrates the basic idea of telescoping very nicely, and the idea of “creative telescoping” turns out to be extremely useful in the study of series — though it is well beyond the scope of this course.
More generally, if we can write
\[ a_n = b_n - b_{n+1} \]
for some other known sequence \( b_n \), then the series telescopes and we can compute partial sums using
\[
\sum_{k=1}^{N} a_k = \sum_{k=1}^{N} (b_k - b_{k+1})
\]
\[
= \sum_{k=1}^{N} b_k - \sum_{k=1}^{N} b_{k+1}
\]
\[
= b_1 - b_{N+1}.
\]
and hence
\[
\sum_{k=1}^{\infty} a_n = b_1 - \lim_{N \to \infty} b_{N+1}
\]
provided this limit exists. Often \( \lim_{N \to \infty} b_{N+1} = 0 \) and then \( \sum_{k=1}^{\infty} a_n = b_1 \). But this does not always happen. Here is an example.

**Example 3.2.7 (A Divergent Telescoping Series)**

In this example, we are going to study the series \( \sum_{n=1}^{\infty} \log \left(1 + \frac{1}{n}\right) \). Let’s start by just writing out the first few terms.
\[
\sum_{n=1}^{\infty} \log \left(1 + \frac{1}{n}\right) = \log \left(1 + 1\right) + \log \left(1 + \frac{1}{2}\right) + \log \left(1 + \frac{1}{3}\right) + \log \left(1 + \frac{1}{4}\right) + \cdots
\]
\[
= \log(2) + \log \left(\frac{3}{2}\right) + \log \left(\frac{4}{3}\right) + \log \left(\frac{5}{4}\right) + \cdots
\]
This is pretty suggestive since
\[
\log(2) + \log \left(\frac{3}{2}\right) + \log \left(\frac{4}{3}\right) + \log \left(\frac{5}{4}\right) = \log \left(2 \times \frac{3}{2} \times \frac{4}{3} \times \frac{5}{4}\right) = \log(5)
\]
So let’s try using this idea to compute the partial sum \( S_N \):
\[
S_N = \sum_{n=1}^{N} \log \left(1 + \frac{1}{n}\right)
\]
\[
= \log \left(1 + \frac{1}{1}\right) + \log \left(1 + \frac{1}{2}\right) + \log \left(1 + \frac{1}{3}\right) + \cdots + \log \left(1 + \frac{1}{N-1}\right) + \log \left(1 + \frac{1}{N}\right)
\]
\[
= \log(2) + \log \left(\frac{3}{2}\right) + \log \left(\frac{4}{3}\right) + \cdots + \log \left(\frac{N}{N-1}\right) + \log \left(\frac{N+1}{N}\right)
\]
\[
= \log \left(2 \times \frac{3}{2} \times \frac{4}{3} \times \cdots \times \frac{N}{N-1} \times \frac{N+1}{N}\right)
\]
\[
= \log(N+1)
\]
Uh oh!

$$\lim_{N \to \infty} S_N = \lim_{N \to \infty} \log(N + 1) = +\infty$$

This telescoping series diverges! There is an important lesson here. Telescoping series can diverge. They do not always converge to $b_1$.

As was the case for limits, differentiation and antidifferentiation, we can compute more complicated series in terms of simpler ones by understanding how series interact with the usual operations of arithmetic. It is, perhaps, not so surprising that there are simple rules for addition and subtraction of series and for multiplication of a series by a constant. Unfortunately there are no simple general rules for computing products or ratios of series.

**Theorem 3.2.8 (Arithmetic of series).**

Let $A$, $B$ and $C$ be real numbers and let the two series $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ converge to $A$ and $B$ respectively. That is, assume that

$$\sum_{n=1}^{\infty} a_n = A \qquad \sum_{n=1}^{\infty} b_n = B$$

Then the following hold.

(a) $\sum_{n=1}^{\infty} [a_n + b_n] = A + B$ and $\sum_{n=1}^{\infty} [a_n - b_n] = A - B$

(b) $\sum_{n=1}^{\infty} Ca_n = CA$.

**Example 3.2.9**

As a simple example of how we use the arithmetic of series Theorem 3.2.8, consider

$$\sum_{n=1}^{\infty} \left[ \frac{1}{7^n} + \frac{2}{n(n + 1)} \right]$$

We recognize that we know how to compute parts of this sum. We know that

$$\sum_{n=1}^{\infty} \frac{1}{7^n} = \frac{1/7}{1 - 1/7} = \frac{1}{6}$$

because it is a geometric series (Example 3.2.4) with first term $a = \frac{1}{7}$ and ratio $r = \frac{1}{7}$. And we know that

$$\sum_{n=1}^{\infty} \frac{1}{n(n + 1)} = 1$$
by Example 3.2.6. We can now use Theorem 3.2.8 to build the specified “complicated”
series out of these two “simple” pieces.

\[
\sum_{n=1}^{\infty} \left[ \frac{1}{7^n} + \frac{2}{n(n+1)} \right] = \sum_{n=1}^{\infty} \frac{1}{7^n} + \sum_{n=1}^{\infty} \frac{2}{n(n+1)} \\
= \sum_{n=1}^{\infty} \frac{1}{7^n} + 2 \sum_{n=1}^{\infty} \frac{1}{n(n+1)} \\
= \frac{1}{6} + 2 \cdot 1 = \frac{13}{6}
\]

by Theorem 3.2.8.a

by Theorem 3.2.8.b

\[ \sum_{n=1}^{\infty} \left[ \frac{1}{7^n} + \frac{2}{n(n+1)} \right] = \sum_{n=1}^{\infty} \frac{1}{7^n} + \sum_{n=1}^{\infty} \frac{2}{n(n+1)} \]

3.3 Convergence Tests

It is very common to encounter series for which it is difficult, or even virtually impos-
ible, to determine the sum exactly. Often you try to evaluate the sum approximately by
truncating it, i.e. having the index run only up to some finite \( N \), rather than infinity. But
there is no point in doing so if the series diverges. So you like to at least know if the
series converges or diverges. Furthermore you would also like to know what error is in-
troduced when you approximate \( \sum_{n=1}^{\infty} a_n \) by the “truncated series” \( \sum_{n=1}^{N} a_n \). That’s called
the truncation error. There are a number of “convergence tests” to help you with this.

3.3.1 The Divergence Test

Our first test is very easy to apply, but it is also rarely useful. It just allows us to quickly
reject some “trivially divergent” series. It is based on the observation that

- by definition, a series \( \sum_{n=1}^{\infty} a_n \) converges to \( S \) when the partial sums \( S_N = \sum_{n=1}^{N} a_n \)
  converge to \( S \).

- Then, as \( N \to \infty \), we have \( S_N \to S \) and, because \( N-1 \to \infty \) too, we also have
  \( S_{N-1} \to S \).

- So \( a_N = S_N - S_{N-1} \to S - S = 0 \).

**Theorem 3.3.1 (Divergence Test).**

If the sequence \( \{a_n\}_{n=1}^{\infty} \) fails to converge to zero as \( n \to \infty \), then the series \( \sum_{n=1}^{\infty} a_n \)
diverges.
Example 3.3.2

Let \( a_n = \frac{n}{n+1} \). Then

\[
\lim_{n \to \infty} a_n = \lim_{n \to \infty} \frac{n}{n+1} = \lim_{n \to \infty} \frac{1}{1 + \frac{1}{n}} = 1 \neq 0
\]

So the series \( \sum_{n=1}^{\infty} \frac{n}{n+1} \) diverges.

Example 3.3.2

Warning 3.3.3.

The divergence test is a “one way test”. It tells us that if \( \lim_{n \to \infty} a_n \) is nonzero, or fails to exist, then the series \( \sum_{n=1}^{\infty} a_n \) diverges. But it tells us absolutely nothing when \( \lim_{n \to \infty} a_n = 0 \). In particular, it is perfectly possible for a series \( \sum_{n=1}^{\infty} a_n \) to diverge even though \( \lim_{n \to \infty} a_n = 0 \). An example is \( \sum_{n=1}^{\infty} \frac{1}{n} \). We’ll show in Example 3.3.5, below, that it diverges.
3.3.2 The Integral Test

**Theorem 3.3.4 (The Integral Test).**

Let $N_0$ be any natural number. If $f(x)$ is a function which is defined and continuous for all $x \geq N_0$ and which obeys

(i) $f(x) \geq 0$ for all $x \geq N_0$ and
(ii) $f(x)$ decreases as $x$ increases and
(iii) $f(n) = a_n$ for all $n \geq N_0$.

Then

$$
\sum_{n=1}^{\infty} a_n \text{ converges } \iff \int_{N_0}^{\infty} f(x) \, dx \text{ converges}
$$

Furthermore, when the series converges, the truncation error

$$
\left| \sum_{n=1}^{\infty} a_n - \sum_{n=1}^{N} a_n \right| \leq \int_{N}^{\infty} f(x) \, dx \quad \text{for all } N \geq N_0
$$

**Proof.** Let $I$ be any fixed integer with $I > N_0$. Then

- $\sum_{n=1}^{\infty} a_n$ converges if and only if $\sum_{n=1}^{\infty} a_n$ converges — removing a fixed finite number of terms from a series cannot impact whether or not it converges.

- Since $a_n \geq 0$ for all $n \geq I > N_0$, the sequence of partial sums $s_\ell = \sum_{n=1}^{\ell} a_n$ obeys $s_{\ell+1} = s_\ell + a_{n+1} \geq s_\ell$. That is, $s_\ell$ increases as $\ell$ increases.

- So $\{s_\ell\}$ must either converge to some finite number or increase to infinity. That is, either $\sum_{n=1}^{\infty} a_n$ converges to a finite number or it is $+\infty$.

Look at the figure above. The shaded area in the figure is $\sum_{n=1}^{\infty} a_n$ because
• the first shaded rectangle has height \( a_I \) and width 1, and hence area \( a_I \) and
• the second shaded rectangle has height \( a_{I+1} \) and width 1, and hence area \( a_{I+1} \), and

so on

This shaded area is smaller than the area under the curve \( y = f(x) \) for \( I - 1 \leq x < \infty \). So

\[
\sum_{n=I}^{\infty} a_n \leq \int_{I-1}^{\infty} f(x) \, dx
\]  

(3.3.1)

and, if the integral is finite, the sum \( \sum_{n=I}^{\infty} a_n \) is finite too.

For the “divergence case” look at the figure above. The (new) shaded area in the figure is again \( \sum_{n=I}^{\infty} a_n \) because

• the first shaded rectangle has height \( a_I \) and width 1, and hence area \( a_I \) and
• the second shaded rectangle has height \( a_{I+1} \) and width 1, and hence area \( a_{I+1} \), and

so on

This time the shaded area is larger than the area under the curve \( y = f(x) \) for \( I \leq x < \infty \). So

\[
\sum_{n=I}^{\infty} a_n \geq \int_{I}^{\infty} f(x) \, dx
\]

and, if the integral is infinite, the sum \( \sum_{n=I}^{\infty} a_n \) is infinite too.

Finally, the bound on the truncation error is just the special case of (3.3.1) with \( I = N + 1 \):

\[
\sum_{n=1}^{\infty} a_n - \sum_{n=1}^{N} a_n = \sum_{n=N+1}^{\infty} a_n \leq \int_{N}^{\infty} f(x) \, dx
\]

Example 3.3.5 (The p test: \( \sum_{n=1}^{\infty} \frac{1}{n^p} \))

Let \( p > 0 \). We’ll now use the integral test to determine whether or not the series \( \sum_{n=1}^{\infty} \frac{1}{n^p} \) (which is sometimes called the p–series) converges.

• To do so, we need a function \( f(x) \) that obeys \( f(n) = a_n = \frac{1}{n^p} \) for all \( n \) bigger than some \( N_0 \). Certainly \( f(x) = \frac{1}{x^p} \) obeys \( f(n) = \frac{1}{n^p} \) for all \( n \geq 1 \). So let’s pick this \( f \) and try \( N_0 = 1 \). (We can always increase \( N_0 \) later if we need to.)
This function also obeys the other two conditions of Theorem 3.3.4:

(i) \( f(x) > 0 \) for all \( x \geq N_0 = 1 \) and
(ii) \( f(x) \) decreases as \( x \) increases because \( f'(x) = -p \frac{1}{x^{p+1}} < 0 \) for all \( x \geq N_0 = 1 \).

So the integral test tells us that the series \( \sum_{n=1}^{\infty} \frac{1}{n^p} \) converges if and only if the integral \( \int_1^{\infty} \frac{dx}{x^p} \) converges.

We have already seen, in Example 1.12.8, that the integral \( \int_1^{\infty} \frac{dx}{x^p} \) converges if and only if \( p > 1 \).

So we conclude that \( \sum_{n=1}^{\infty} \frac{1}{n^p} \) converges if and only if \( p > 1 \). This is sometimes called the \( p \)-test.

In particular, the series \( \sum_{n=1}^{\infty} \frac{1}{n} \), which is called the harmonic series, has \( p = 1 \) and so diverges. As we add more and more terms of this series together, the terms we add, namely \( \frac{1}{n} \), get smaller and smaller and tend to zero, but the rate at which they tend to zero is small enough that the full sum is still infinite.

On the other hand, the series \( \sum_{n=1}^{\infty} \frac{1}{n^{1.000001}} \) has \( p = 1.000001 > 1 \) and so converges. This time as we add more and more terms of this series together, the terms we add, namely \( \frac{1}{n^{1.000001}} \), tend to zero (just) fast enough that the full sum is finite. Mind you, for this example, the convergence takes place very slowly — you have to take a huge number of terms to get a decent approximation to the full sum. If we approximate \( \sum_{n=1}^{\infty} \frac{1}{n^{1.000001}} \) by the truncated series \( \sum_{n=1}^{N} \frac{1}{n^{1.000001}} \), we make an error of at most

\[
\int_N^{\infty} \frac{dx}{x^{1.000001}} = \lim_{R \to \infty} \int_N^R \frac{dx}{x^{1.000001}} = \lim_{R \to \infty} \frac{1}{R^{0.000001}} - \frac{1}{N^{0.000001}} \approx \frac{10^6}{N^{0.000001}}
\]

This does tend to zero as \( N \to \infty \), but really slowly.

Let \( p > 0 \). We’ll now use the integral test to determine whether or not the series \( \sum_{n=2}^{\infty} \frac{1}{n(\log n)^p} \) converges.

As in the last example, we start by choosing a function that obeys \( f(n) = a_n = \frac{1}{n(\log n)^p} \) for all \( n \) bigger than some \( N_0 \). Certainly \( f(x) = \frac{1}{x(\log x)^p} \) obeys \( f(n) = \frac{1}{n(\log n)^p} \) for all \( n \geq 2 \). So let’s use that \( f \) and try \( N_0 = 2 \).

Now let’s check the other two conditions of Theorem 3.3.4:

(i) Both \( x \) and \( \log x \) are positive for all \( x > 1 \), so \( f(x) > 0 \) for all \( x \geq N_0 = 2 \).
(ii) As \( x \) increases both \( x \) and \( \log x \) increase and so \( x(\log x)^p \) increases and \( f(x) \) decreases.
• So the integral test tells us that the series \( \sum_{n=2}^{\infty} \frac{1}{n \log n} \) converges if and only if the integral \( \int_{2}^{\infty} \frac{dx}{x \log x} \) converges.

• To test the convergence of the integral, we make the substitution \( u = \log x \), \( du = \frac{dx}{x} \).

\[
\int_{2}^{R} \frac{dx}{x \log x} = \int_{\log 2}^{\log R} \frac{du}{u^p}
\]

We already know that the integral \( \int_{1}^{\infty} \frac{du}{u^p} \), and hence the integral \( \int_{2}^{R} \frac{dx}{x \log x} \), converges if and only if \( p > 1 \).

So we conclude that \( \sum_{n=1}^{\infty} \frac{1}{n \log n} \) converges if and only if \( p > 1 \).

### 3.3.3 The Comparison Test

Our next convergence test is the comparison test. It is much like the comparison test for improper integrals (see Theorem 1.12.17) and is true for much the same reasons.

#### Theorem 3.3.7 (The Comparison Test)

Let \( N_0 \) be a natural number and let \( K > 0 \).

(a) If \( |a_n| \leq Kc_n \) for all \( n \geq N_0 \) and \( \sum_{n=0}^{\infty} c_n \) converges, then \( \sum_{n=0}^{\infty} a_n \) converges.

(b) If \( a_n \geq Kd_n \geq 0 \) for all \( n \geq N_0 \) and \( \sum_{n=0}^{\infty} d_n \) diverges, then \( \sum_{n=0}^{\infty} a_n \) diverges.

“Proof”. We will not prove this theorem. We’ll just observe that it is very reasonable. That’s why there are quotation marks around “Proof”.

(a) If \( \sum_{n=0}^{\infty} c_n \) converges to a finite number and if the terms in \( \sum_{n=0}^{\infty} a_n \) are smaller than the terms in \( \sum_{n=0}^{\infty} c_n \), then it is no surprise that \( \sum_{n=0}^{\infty} a_n \) converges too.

(b) If \( \sum_{n=0}^{\infty} d_n \) diverges (i.e. adds up to \( \infty \)) and if the terms in \( \sum_{n=0}^{\infty} a_n \) are larger than the terms in \( \sum_{n=0}^{\infty} d_n \), then of course \( \sum_{n=0}^{\infty} a_n \) adds up to \( \infty \), and so diverges, too.

The comparison test for series is also used in much the same way as is the comparison test for improper integrals.
Example 3.3.8 \( \sum_{n=1}^{\infty} \frac{1}{n^2 + 2n + 3} \)

We could determine whether or not the series \( \sum_{n=1}^{\infty} \frac{1}{n^2 + 2n + 3} \) converges by applying the integral test. But it is not worth the effort. Whether or not any series converges is determined by the behaviour of the summand for very large \( n \). So the first step in tackling such a problem is to develop some intuition about the behaviour of \( a_n \) when \( n \) is very large.

- **Step 1: Develop intuition.** In this case, when \( n \) is very large\(^{13} \), \( n^2 \gg 2n \gg 3 \) so that \( \frac{1}{n^2 + 2n + 3} \approx \frac{1}{n^2} \). We already know, from Example 3.3.5, that \( \sum_{n=1}^{\infty} \frac{1}{n^p} \) converges if and only if \( p > 1 \). So \( \sum_{n=1}^{\infty} \frac{1}{n^2} \), which has \( p = 2 \), converges, and we would expect that \( \sum_{n=1}^{\infty} \frac{1}{n^2 + 2n + 3} \) converges too.

- **Step 2: Verify intuition.** We can use the comparison test to confirm that this is indeed the case. For any \( n \geq 1 \), \( n^2 + 2n + 3 > n^2 \), so that \( \frac{1}{n^2 + 2n + 3} \leq \frac{1}{n^2} \). So the comparison test, Theorem 3.3.7, with \( a_n = \frac{1}{n^2 + 2n + 3} \) and \( c_n = \frac{1}{n^2} \), tells us that \( \sum_{n=1}^{\infty} \frac{1}{n^2 + 2n + 3} \) converges.

Example 3.3.9 \( \sum_{n=1}^{\infty} \frac{n + \cos n}{n^3 - 1/3} \)

As in the previous example, the first step is to develop some intuition about the behaviour of \( a_n \) when \( n \) is very large.

- **Step 1: Develop intuition.** When \( n \) is very large,
  
  - \( n \gg \mid \cos n \mid \) so that the numerator \( n + \cos n \approx n \) and
  - \( n^3 \gg 1/3 \) so that the denominator \( n^3 - 1/3 \approx n^3 \).

  So when \( n \) is very large
  \[
a_n = \frac{n + \cos n}{n^3 - 1/3} \approx \frac{n}{n^3} = \frac{1}{n^2}
  \]

  We already know from Example 3.3.5, with \( p = 2 \), that \( \sum_{n=1}^{\infty} \frac{1}{n^p} \) converges, so we would expect that \( \sum_{n=1}^{\infty} \frac{n + \cos n}{n^3 - 1/3} \) converges too.

- **Step 2: Verify intuition.** We can use the comparison test to confirm that this is indeed the case. To do so we need to find a constant \( K \) such that \( \mid a_n \mid = \frac{\mid n + \cos n \mid}{n^3 - 1/3} = \frac{n + \cos n}{n^3 - 1/3} \) is smaller than \( \frac{K}{n^2} \) for all \( n \). A good way to do that is to factor the dominant term (in

\[^{13}\text{The symbol “\( \gg \)” means “much larger than”}\]
this case $n$) out of the numerator and also factor the dominant term (in this case $n^3$)
out of the denominator.

$$a_n = \frac{n + \cos n}{n^3 - 1/3} = \frac{n}{n^3} \frac{1 + \cos n}{1 - \frac{1}{3n^3}} = \frac{1}{n^2} \frac{1 + \cos n}{1 - \frac{1}{3n^3}}$$

So now we need to find a constant $K$ such that $\frac{1 + (\cos n)/n}{1 - 1/3n^3}$ is smaller than $K$ for all $n \geq 1$.

- First consider the numerator $1 + (\cos n) \frac{1}{n}$. For all $n \geq 1$
  - $\frac{1}{n} \leq 1$ and
  - $|\cos n| \leq 1$

So the numerator $1 + (\cos n) \frac{1}{n}$ is always smaller than $1 + (1) \frac{1}{1} = 2$.

- Next consider the denominator $1 - 1/3n^3$.
  - When $n = 1$, $\frac{1}{1 - 1/3n^3} = \frac{1}{2/3} = \frac{3}{2}$.
  - As $n$ increases, $3n^3$ increases, so $1/3n^3$ decreases towards 0, so $1 - 1/3n^3$ increases towards 1 and $\frac{1}{1 - 1/3n^3}$ decreases towards 1.
  - Consequently $\frac{1}{1 - 1/3n^3}$ is never bigger than $\frac{3}{2}$

- As the numerator $1 + (\cos n) \frac{1}{n}$ is always smaller than 2 and $\frac{1}{1 - 1/3n^3}$ is always smaller than $\frac{3}{2}$, the fraction

$$\frac{1 + \cos n}{1 - \frac{1}{3n^3}} \leq 2 \left( \frac{3}{2} \right) = 3$$

We now know that

$$|a_n| = \frac{1}{n^2} \frac{1 + 2/n}{1 - 1/3n^3} \leq \frac{3}{n^2}$$

and the comparison test tells us that $\sum_{n=1}^{\infty} \frac{n + \cos n}{n^3 - 1/3}$ converges.

---

The last example was actually a relatively simple application of the comparison theorem — finding a suitable constant $K$ can be really tedious. Fortunately, there is a variant of the comparison test that completely eliminates the need to find $K$. 

---

Example 3.3.9

---

277
Theorem 3.3.10 (Limit Comparison Theorem).

Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be two series with $b_n > 0$ for all $n$. Assume that

$$\lim_{n \to \infty} \frac{a_n}{b_n} = L$$

exists.

(a) If $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ converges too.

(b) If $L \neq 0$ and $\sum_{n=1}^{\infty} b_n$ diverges, then $\sum_{n=1}^{\infty} a_n$ diverges too.

In particular, if $L \neq 0$, then $\sum_{n=1}^{\infty} a_n$ converges if and only if $\sum_{n=1}^{\infty} b_n$ converges.

Proof. We will not prove this theorem, but we will explain the idea behind the proof.

(a) Because we are told that $\lim_{n \to \infty} \frac{a_n}{b_n} = L$, we know that,

- when $n$ is large, $\frac{a_n}{b_n}$ is very close to $L$, so that $\left| \frac{a_n}{b_n} \right|$ is very close to $|L|$.

- In particular, there is some natural number $N$ so that $\left| \frac{a_n}{b_n} \right| \leq |L| + 1$, and hence

$|a_n| \leq Kb_n$ with $K = |L| + 1$, for all $n \geq N$.

- The comparison Theorem 3.3.7 now implies that $\sum_{n=1}^{\infty} a_n$ converges.

(b) Let’s suppose that $L > 0$. (If $L < 0$, just replace $a_n$ with $-a_n$.) Because we are told that $\lim_{n \to \infty} \frac{a_n}{b_n} = L$, we know that,

- when $n$ is large, $\frac{a_n}{b_n}$ is very close to $L$.

- In particular, there is some natural number $N$ so that $\frac{a_n}{b_n} \geq \frac{L}{2}$, and hence

$|a_n| \geq Kb_n$ with $K = \frac{L}{2} > 0$, for all $n \geq N$.

- The comparison Theorem 3.3.7 now implies that $\sum_{n=1}^{\infty} a_n$ diverges.

Example 3.3.11 $\left( \sum_{n=1}^{\infty} \frac{\sqrt{n+1}}{n^2 - 2n + 3} \right)$

Set $a_n = \frac{\sqrt{n+1}}{n^2 - 2n + 3}$. We first try to develop some intuition about the behaviour of $a_n$ for large $n$ and then we confirm that our intuition was correct.

- Step 1: Develop intuition. When $n \gg 1$, the numerator $\sqrt{n+1} \approx \sqrt{n}$, and the denominator $n^2 - 2n + 3 \approx n^2$ so that $a_n \approx \frac{\sqrt{n}}{n^2} = \frac{1}{n^{3/2}}$ and it looks like our series should converge by Example 3.3.5 with $p = \frac{3}{2}$. 

278
• **Step 2: Verify intuition.** To confirm our intuition we set $b_n = \frac{1}{n^{3/2}}$ and compute the limit

$$\lim_{n \to \infty} \frac{a_n}{b_n} = \lim_{n \to \infty} \frac{n^{3/2} \sqrt{n + 1}}{n^{3/2} - 2n + 3} = \lim_{n \to \infty} \frac{n^{3/2} \sqrt{n + 1}}{n^{3/2} - 2n + 3}$$

Again it is a good idea to factor the dominant term out of the numerator and the dominant term out of the denominator.

$$\lim_{n \to \infty} \frac{a_n}{b_n} = \lim_{n \to \infty} \frac{n^{2/3} \sqrt{1 + 1/n}}{n^2 (1 - 2/n + 3/n^2)} = \lim_{n \to \infty} \frac{\sqrt{1 + 1/n}}{1 - 2/n + 3/n^2} = 1$$

We already know that the series $\sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} \frac{1}{n^{3/2}}$ converges by Example 3.3.5 with $p = \frac{3}{2}$. So our series converges by the limit comparison test, Theorem 3.3.10.

### 3.3.4 The Alternating Series Test

When the signs of successive terms in a series alternate between $+$ and $-$, like for example in $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$, the series is called an *alternating series*. More generally, the series

$$a_1 - a_2 + a_3 - a_4 + \cdots = \sum_{n=1}^{\infty} (-1)^{n-1} a_n$$

is alternating if every $a_n \geq 0$. Often (but not always) the terms in alternating series get successively smaller. That is, then $a_1 \geq a_2 \geq a_3 \geq \cdots$. In this case:

- The first partial sum is $S_1 = a_1$.
- The second partial sum, $S_2 = a_1 - a_2$, is smaller than $S_1$ by $a_2$.
- The third partial sum, $S_3 = S_2 + a_3$, is bigger than $S_2$ by $a_3$, but because $a_3 \leq a_2$, $S_3$ remains smaller than $S_1$. See the figure below.
- The fourth partial sum, $S_4 = S_3 - a_4$, is smaller than $S_3$ by $a_4$, but because $a_4 \leq a_3$, $S_4$ remains bigger than $S_2$. Again, see the figure below.
- And so on.

So the successive partial sums oscillate, but with ever decreasing amplitude. If, in addition, $a_n$ tends to $0$ as $n$ tends to $\infty$, the amplitude of oscillation tends to zero and the sequence $S_1, S_2, S_3, \cdots$ converges. This is illustrated in the figure.
Here is a convergence test for alternating series that exploits this structure, and that is really easy to apply.

**Theorem 3.3.12 (Alternating Series Test).**

Let \( \{a_n\}_{n=1}^{\infty} \) be a sequence of real numbers that obeys

(i) \( a_n \geq 0 \) for all \( n \geq 1 \) and
(ii) \( a_{n+1} \leq a_n \) for all \( n \geq 1 \) (i.e. the sequence is monotone decreasing) and
(iii) \( \lim_{n \to \infty} a_n = 0. \)

Then

\[
a_1 - a_2 + a_3 - a_4 + \cdots = \sum_{n=1}^{\infty} (-1)^{n-1} a_n = S
\]

converges and, for each natural number \( N \), \( S - S_N \) is between 0 and (the first dropped term) \( (-1)^N a_{N+1} \).

"Proof". We are not going to give a complete proof. We shall fix any natural number \( N \) and concentrate on the last statement, which gives a bound on the truncation error (which is the error introduced when you approximate the full series by the partial sum \( S_N \))

\[
E_N = S - S_N = \sum_{n=N+1}^{\infty} (-1)^{n-1} a_n = (-1)^N \left[ a_{N+1} - a_{N+2} + a_{N+3} - a_{N+4} + \cdots \right]
\]

This is of course another series. We’re going to study the partial sums

\[
S_{N,\ell} = \sum_{n=N+1}^{\ell} (-1)^{n-1} a_n
\]

for that series.
• If \( \ell > N + 1 \), with \( \ell - N \) even,

\[
(-1)^N S_{N,\ell} = \sum_{i=0}^{\frac{\ell - 1}{2}} (a_{N+i} - a_{N+i+1}) + (a_{N+i+2} - a_{N+i+3}) + \cdots + (a_{N+i+\frac{\ell - 1}{2}} - a_{N+i+\frac{\ell}{2}}) \geq 0 \quad \text{and}
\]

\[
(-1)^N S_{N,\ell+1} = \frac{1}{S_{N,\ell}} - \frac{a_{\ell+1}}{a_{\ell+2}} \geq 0
\]

This tells us that \((-1)^N S_{N,\ell} \geq 0\) for all \( \ell > N + 1 \).

• Similarly, if \( \ell > N + 1 \), with \( \ell - N \) odd,

\[
(-1)^N S_{N,\ell} = a_{N+1} - (a_{N+2} - a_{N+3}) - (a_{N+4} - a_{N+5}) - \cdots - (a_{\ell-1} - a_{\ell}) \leq a_{N+1} \quad \text{and}
\]

\[
(-1)^N S_{N,\ell+1} = \frac{S_{N,\ell}}{a_{\ell+1}} - \frac{1}{a_{\ell+2}} \leq a_{N+1}
\]

This tells us that \((-1)^N S_{N,\ell} \leq a_{N+1}\) for all \( \ell > N + 1 \).

So we now know that \( S_{N,\ell} \) lies between its first term, \((-1)^N a_{N+1}\), and 0 for all \( \ell > N + 1 \). While we are not going to prove it, this implies that, since \( a_{N+1} \to 0 \) as \( N \to \infty \), the series \( S \) converges and that

\[
S - S_N = \lim_{\ell \to \infty} S_{N,\ell}
\]

lies between \((-1)^N a_{N+1}\), and 0.

---

**Example 3.3.13**

We have already seen, in Example 3.3.5, that the harmonic series \( \sum_{n=1}^{\infty} \frac{1}{n} \) diverges. On the other hand, the series \( \sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} \) converges by the alternating series test with \( a_n = \frac{1}{n} \).

Note that

(i) \( a_n = \frac{1}{n} \geq 0 \) for all \( n \geq 1 \), so that \( \sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} \) really is an alternating series, and

(ii) \( a_n = \frac{1}{n} \) decreases as \( n \) increases, and

(iii) \( \lim_{n \to \infty} a_n = \lim_{n \to \infty} \frac{1}{n} = 0 \).

so that all of the hypotheses of the alternating series test, i.e. of Theorem 3.3.12, are satisfied.

---

**Example 3.3.14 (e)**

You may already know that \( e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \). In any event, we shall prove this in Example 3.6.1, below. In particular

\[
\frac{1}{e} = e^{-1} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \cdots
\]

is an alternating series and satisfies all of the conditions of the alternating series test, Theorem 3.3.12a:
So the alternating series test guarantees that, if we approximate, for example,
\[
\frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} + \frac{1}{8!} - \frac{1}{9!}
\]
then the error in this approximation lies between 0 and the next term in the series, which is \(\frac{1}{10!}\). That is
\[
\frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} + \frac{1}{8!} - \frac{1}{9!} \leq \frac{1}{e} \leq \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} + \frac{1}{8!} - \frac{1}{9!} + \frac{1}{10!}
\]
so that
\[
\frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} + \frac{1}{8!} - \frac{1}{9!} - \frac{1}{10!} \leq e \leq \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} + \frac{1}{8!} - \frac{1}{9!} - \frac{1}{10!}
\]
which, to seven decimal places says
\[
2.7182816 \leq e \leq 2.7182837
\]
(To seven decimal places \(e = 2.7182818\).)

**Example 3.3.14**

3.3.5 The Ratio Test

**Theorem 3.3.15 (Ratio Test).**

Let \(N\) be any positive integer and assume that \(a_n \neq 0\) for all \(n \geq N\).

(a) If \(\lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = L < 1\), then \(\sum_{n=1}^{\infty} a_n\) converges.

(b) If \(\lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = L > 1\), or \(\lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = +\infty\), then \(\sum_{n=1}^{\infty} a_n\) diverges.

**Proof:** (a) Pick any number \(R\) obeying \(L < R < 1\). We are assuming that \(\left| \frac{a_{n+1}}{a_n} \right|\) approaches \(L\) as \(n \to \infty\). In particular there must be some natural number \(M\) so that \(\left| \frac{a_{n+1}}{a_n} \right| \leq R\) for all \(n \geq M\). So \(|a_{n+1}| \leq R|a_n|\) for all \(n \geq M\). In particular
\[
|a_{M+1}| \leq R|a_M|
\]
\[
|a_{M+2}| \leq R|a_{M+1}| \leq R^2 |a_M|
\]
\[
|a_{M+3}| \leq R|a_{M+2}| \leq R^3 |a_M|
\]
\[
\vdots
\]
\[
|a_{M+\ell}| \leq R^\ell |a_M|
\]
for all \( \ell \geq 0 \). The series \( \sum_{\ell=0}^{\infty} R^\ell |a_M| \) is a geometric series with ratio \( R \) smaller than one in magnitude and so converges. Consequently, by the comparison test with \( a_n \) replaced by \( A_\ell = a_{n+\ell} \) and \( c_n \) replaced by \( C_\ell = R^\ell |a_M| \), the series \( \sum_{\ell=1}^{\infty} a_{M+\ell} = \sum_{n=M+1}^{\infty} a_n \) converges. So the series \( \sum_{n=1}^{\infty} a_n \) converges too.

(b) We are assuming that \( \left| \frac{a_{n+1}}{a_n} \right| \) approaches \( L > 1 \) as \( n \to \infty \). In particular there must be some natural number \( M > N \) so that \( \left| \frac{a_{n+1}}{a_n} \right| \geq 1 \) for all \( n \geq M \). So \( |a_{n+1}| \geq |a_n| \) for all \( n \geq M \). That is, \( |a_n| \) increases as \( n \) increases as long as \( n \geq M \). So \( |a_n| \geq |a_M| \) for all \( n \geq M \) and \( a_n \) cannot converge to zero as \( n \to \infty \). So the series diverges by the divergence test.

**Warning 3.3.16.**

Beware that the ratio test provides absolutely no conclusion about the convergence or divergence of the series \( \sum_{n=1}^{\infty} a_n \) if \( \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = 1 \). See Example 3.3.19, below.

---

**Example 3.3.17 \( \sum_{n=0}^{\infty} anx^{n-1} \)**

Fix any two nonzero real numbers \( a \) and \( x \). We have already seen in Example 3.2.4 — we have just renamed \( r \) to \( x \) — that the geometric series \( \sum_{n=0}^{\infty} ax^n \) converges when \( |x| < 1 \) and diverges when \( |x| \geq 1 \). We are now going to consider a new series, constructed by differentiating each term in the geometric series \( \sum_{n=0}^{\infty} ax^n \). This new series is

\[
\sum_{n=0}^{\infty} a_n \text{ with } a_n = a n x^{n-1}
\]

Let’s apply the ratio test.

\[
\left| \frac{a_{n+1}}{a_n} \right| = \frac{a (n+1) x^n}{a n x^{n-1}} = \frac{n+1}{n} |x| = \left( 1 + \frac{1}{n} \right) |x| \to L = |x| \text{ as } n \to \infty
\]

The ratio test now tells us that the series \( \sum_{n=0}^{\infty} an x^{n-1} \) converges if \( |x| < 1 \) and diverges if \( |x| > 1 \). It says nothing about the cases \( x = \pm 1 \). But in both of those cases \( a_n = a n (\pm 1)^n \) does not converge to zero as \( n \to \infty \) and the series diverges by the divergence test.

---

We shall see later, in Theorem 3.5.11, that the function \( \sum_{n=0}^{\infty} anx^{n-1} \) is indeed the derivative of the function \( \sum_{n=0}^{\infty} ax^n \).
Example 3.3.18 \( \left( \sum_{n=0}^{\infty} \frac{a}{n+1} X^{n+1} \right) \)

Once again, fix any two nonzero real numbers \( a \) and \( X \). We again start with the geometric series \( \sum_{n=0}^{\infty} a X^n \) but this time we construct a new series by integrating each term, \( a X^n \), from \( x = 0 \) to \( x = X \) giving \( \sum_{n=0}^{\infty} \frac{a}{n+1} X^{n+1} \). The resulting new series is

\[
\sum_{n=0}^{\infty} a_n \quad \text{with} \quad a_n = \frac{a}{n+1} X^{n+1}
\]

To apply the ratio test we need to compute

\[
\left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{\frac{a}{n+2} X^{n+2}}{\frac{a}{n+1} X^{n+1}} \right| = \frac{n+1}{n+2} |X| = \frac{1 + \frac{1}{n} |X|}{1 + \frac{2}{n}} \to L = |X| \quad \text{as} \quad n \to \infty
\]

The ratio test now tells us that the series \( \sum_{n=0}^{\infty} \frac{a}{n+1} X^{n+1} \) converges if \( |X| < 1 \) and diverges if \( |X| > 1 \). It says nothing about the cases \( X = \pm 1 \).

If \( X = 1 \), the series reduces to

\[
\sum_{n=0}^{\infty} \frac{a}{n+1} X^{n+1} \bigg|_{X=1} = \sum_{n=0}^{\infty} \frac{a}{n+1} = a \sum_{m=1}^{\infty} \frac{1}{m}
\]

which is just \( a \) times the harmonic series, which we know diverges, by Example 3.3.5.

If \( X = -1 \), the series reduces to

\[
\sum_{n=0}^{\infty} \frac{a}{n+1} X^{n+1} \bigg|_{X=-1} = \sum_{n=0}^{\infty} (-1)^{n+1} \frac{a}{n+1}
\]

which converges by the alternating series test. See Example 3.3.13.

In conclusion, the series \( \sum_{n=0}^{\infty} \frac{a}{n+1} X^{n+1} \) converges if and only if \(-1 \leq X < 1 \).

Example 3.3.19 \( (L = 1) \)

In this example, we are going to see three different series that have \( \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = 1 \). One is going to diverge and the other two are going to converge.

- The first series is the harmonic series

\[
\sum_{n=1}^{\infty} a_n \quad \text{with} \quad a_n = \frac{1}{n}
\]

We have already seen, in Example 3.3.5, that this series diverges. It has

\[
\left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{\frac{1}{n+1}}{\frac{1}{n}} \right| = \frac{n}{n+1} = \frac{1}{1+\frac{1}{n}} \to L = 1 \quad \text{as} \quad n \to \infty
\]

Example 3.3.18

We shall also see later, in Theorem 3.5.11, that the function \( \sum_{n=0}^{\infty} \frac{a}{n+1} X^{n+1} \) is indeed an antiderivative of the function \( \sum_{n=0}^{\infty} a X^n \).
• The second series is the alternating harmonic series

\[ \sum_{n=1}^{\infty} a_n \quad \text{with } a_n = (-1)^{n-1} \frac{1}{n} \]

We have already seen, in Example 3.3.13, that this series converges. But it also has

\[ \left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{(-1)^n \frac{1}{n+1}}{(-1)^{n-1} \frac{1}{n}} \right| = \frac{n}{n+1} = \frac{1}{1 + \frac{1}{n}} \rightarrow L = 1 \quad \text{as } n \rightarrow \infty \]

• The third series is

\[ \sum_{n=1}^{\infty} a_n \quad \text{with } a_n = \frac{1}{n^2} \]

We have already seen, in Example 3.3.5 with \( p = 2 \), that this series converges. But it also has

\[ \left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{\frac{1}{(n+1)^2}}{\frac{1}{n^2}} \right| = \frac{n^2}{(n+1)^2} = \frac{1}{(1 + \frac{1}{n})^2} \rightarrow L = 1 \quad \text{as } n \rightarrow \infty \]

3.3.6 Convergence Test List

We now have half a dozen convergence tests:

• **Divergence Test**
  - works well when the \( n^{\text{th}} \) term in the series fails to converge to zero as \( n \) tends to infinity

• **Alternating Series Test**
  - works well when successive terms in the series alternate in sign
  - don’t forget to check that successive terms decrease in magnitude and tend to zero as \( n \) tends to infinity

• **Integral Test**
  - works well when, if you substitute \( x \) for \( n \) in the \( n^{\text{th}} \) term you get a function, \( f(x) \), that you can integrate
  - don’t forget to check that \( f(x) \geq 0 \) and that \( f(x) \) decreases as \( x \) increases

• **Ratio Test**
  - works well when \( \frac{a_{n+1}}{a_n} \) simplifies enough that you can easily compute \( \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = L \)
– this often happens when $a_n$ contains powers, like $7^n$, or factorials, like $n!$
– don’t forget that $L = 1$ tells you nothing about the convergence/divergence of the series

• **Comparison Test and Limit Comparison Test**
  – works well when, for very large $n$, the $n$th term $a_n$ is approximately the same as a simpler term $b_n$ (see Example 3.3.9) and it is easy to determine whether or not $\sum_{n=1}^{\infty} b_n$ converges
  – don’t forget to check that $b_n \geq 0$
  – usually the Limit Comparison Test is easier to apply than the Comparison Test

### 3.4 Absolute and Conditional Convergence

#### 3.4.1 Definitions

**Definition 3.4.1 (Absolute and conditional convergence).**

(a) A series $\sum_{n=1}^{\infty} a_n$ is said to converge absolutely if the series $\sum_{n=1}^{\infty} |a_n|$ converges.

(b) If $\sum_{n=1}^{\infty} a_n$ converges but $\sum_{n=1}^{\infty} |a_n|$ diverges we say that $\sum_{n=1}^{\infty} a_n$ is conditionally convergent.

**Theorem 3.4.2 (Absolute convergence implies convergence).**

If the series $\sum_{n=1}^{\infty} |a_n|$ converges then the series $\sum_{n=1}^{\infty} a_n$ also converges. That is, absolute convergence implies convergence.

Recall that some of our convergence tests (for example, the integral test) may only be applied to series with positive terms. Theorem 3.4.2 opens up the possibility of applying “positive only” convergence tests to series whose terms are not all positive, by checking for “absolute convergence” rather than for plain “convergence”.

**Example 3.4.3** \(\sum_{n=1}^{\infty} \left(\frac{(-1)^{n-1}}{n}\right)\)

The alternating harmonic series \(\sum_{n=1}^{\infty} \left(\frac{(-1)^{n-1}}{n}\right)\) of Example 3.3.13 converges (by the alternating series test). But the harmonic series \(\sum_{n=1}^{\infty} \frac{1}{n}\) of Example 3.3.5 diverges (by the integral test). So the alternating harmonic series \(\sum_{n=1}^{\infty} \left(\frac{(-1)^{n-1}}{n}\right)\) converges conditionally.
Example 3.4.3

Example 3.4.4 \( \left( \sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} \right) \)

Because the series \( \sum_{n=1}^{\infty} \left| (-1)^{n-1} \frac{1}{n} \right| = \sum_{n=1}^{\infty} \frac{1}{n^2} \) of Example 3.3.5 converges (by the integral test), the series \( \sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} \) converges absolutely, and hence converges.

Example 3.4.4

Example 3.4.5 (random signs)

Imagine flipping a coin infinitely many times. Set \( \sigma_n = +1 \) if the \( n \)th flip comes up heads and \( \sigma_n = -1 \) if the \( n \)th flip comes up tails. The series \( \sum_{n=1}^{\infty} (-1)^{\sigma_n} \frac{1}{n} \) is not in general an alternating series. But we know that the series \( \sum_{n=1}^{\infty} \left| (-1)^{\sigma_n} \frac{1}{n} \right| = \sum_{n=1}^{\infty} \frac{1}{n^2} \) converges. So \( \sum_{n=1}^{\infty} (-1)^{\sigma_n} \frac{1}{n} \) converges absolutely, and hence converges.

Example 3.4.5

3.4.2 Optional — The delicacy of conditionally convergent series

Conditionally convergent series have to be treated with great care. For example, switching the order of the terms in a finite sum does not change its value.

\[ 1 + 2 + 3 + 4 + 5 + 6 = 6 + 3 + 5 + 2 + 4 + 1 \]

The same is true for absolutely convergent series. But it is not true for conditionally convergent series. In fact by reordering any conditionally convergent series, you can make it add up to any number you like, including \( +\infty \) and \(-\infty \). Here is an example that shows why.

Example 3.4.6

In this example, we’ll reorder the conditionally convergent series \( \sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} \) so that it adds up to exactly 1.234. (Of course, the target 1.234 has been chosen at random. It can be replaced by any number you like.) First create two lists of numbers — the first list consisting of the positive terms of the series, in order, and the second consisting of the negative numbers of the series, in order.

\[ 1, \frac{1}{3}, \frac{1}{5}, \frac{1}{7}, \cdots \quad \frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \cdots \]

Note that if we add together the numbers in the second list, we get \(-\frac{1}{2} \left[ 1 + \frac{1}{2} + \frac{1}{3} + \cdots \right] \), which is just \(-\frac{1}{2}\) times the harmonic series. So the numbers in the second list add up to...
\(-\infty\). Also, if we add together the numbers in the first list, we get \(1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \cdots\) which is bigger than \(\frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \frac{1}{8} = \frac{1}{2} [1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots]\). So the numbers in the first list add up to \(+\infty\).

Now we build up our reordered series. Start by moving just enough numbers from the beginning of the first list into the reordered series to get a sum bigger than 1.234.

\[
1 + \frac{1}{3} = 1.3333
\]

Next move just enough numbers from the beginning of the second list into the reordered series to get a number less than 1.234.

\[
1 + \frac{1}{3} - \frac{1}{2} = 0.8333
\]

Next move just enough numbers from the beginning of the remaining part of the first list into the reordered series to get a number bigger than 1.234.

\[
1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} + \frac{1}{9} = 1.2873
\]

Next move just enough numbers from the beginning of the remaining part of the second list into the reordered series to get a number less than 1.234.

\[
1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} + \frac{1}{9} - \frac{1}{4} = 1.0373
\]

Just keep going like this. At the end of each step, the difference between the sum and 1.234 is smaller than the magnitude of the first unused number in the list. Since the numbers in both lists tend to zero as you go farther and farther up the list, this procedure will generate a series whose sum is exactly 1.234. Since in each step we remove at least one number from a list and we alternate between the two lists, the reordered series will contain all of the terms from \(\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n}\), with each term appearing exactly once.

---

### 3.5 Power Series

#### 3.5.1 Definitions

Remember that we set as our goal, in studying sequences and series, the development of machinery which would allow us to answer questions like, “Is \(e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}\)?”. We are now ready to start working on series like \(\sum_{n=0}^{\infty} \frac{x^n}{n!}\). We’ll start with the definition of a power series.
A series of the form
\[
A_0 + A_1(x - c) + A_2(x - c)^2 + A_3(x - c)^3 + \cdots = \sum_{n=0}^{\infty} A_n(x - c)^n
\]
is called a power series in \((x - c)\) or a power series centered about \(c\). The numbers \(A_n\) are called the coefficients of the power series. Often \(c = 0\) and then the series reduces to
\[
A_0 + A_1x + A_2x^2 + A_3x^3 + \cdots = \sum_{n=0}^{\infty} A_n x^n
\]

For example \(\sum_{n=0}^{\infty} \frac{x^n}{n!}\) is the power series with \(c = 0\) and \(A_n = \frac{1}{n!}\). Typically, as in \(\sum_{n=0}^{\infty} \frac{x^n}{n!}\), the coefficients \(A_n\) are given fixed numbers. But the \(x\) in a power series is to be thought of as a variable. So each power series is really a whole family of series — a different series for each value of \(x\). One possible value of \(x\) is \(x = c\) and then the series reduces to
\[
\sum_{n=0}^{\infty} A_n(x - c)^n = \sum_{n=0}^{\infty} A_n (c - c)^n = \sum_{n=0}^{\infty} A_n 0^n = A_0 + 0 + 0 + 0 + \cdots
\]
which trivially converges to \(A_0\).

Let’s apply the ratio test to try and determine which for other values of \(x\) this series also converges. The \(n^{th}\) term in the series \(\sum_{n=0}^{\infty} A_n(x - c)^n\) is \(a_n = A_n(x - c)^n\). So the ratio test tells us to compute
\[
\left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{A_{n+1}(x - c)^{n+1}}{A_n(x - c)^n} \right| = \left| \frac{A_{n+1}}{A_n} \right| |x - c|
\]
Now we are to try and take the limit \(n \to \infty\). There are several possibilities.

- If the limit \(\lim_{n \to \infty} \left| \frac{A_{n+1}}{A_n} \right|\) exists and equals some nonzero value, say \(A\), then the ratio test says that the series \(\sum_{n=0}^{\infty} A_n(x - c)^n\) converges when \(A|x - c| < 1\), i.e. when \(|x - c| < \frac{1}{A}\), and diverges when \(A|x - c| > 1\), i.e. when \(|x - c| > \frac{1}{A}\). This
\[
R = \frac{1}{A} = \left[ \lim_{n \to \infty} \left| \frac{A_{n+1}}{A_n} \right| \right]^{-1}
\]
is called the radius of convergence of the series.

- If the limit \(\lim_{n \to \infty} \left| \frac{A_{n+1}}{A_n} \right|\) exists and equals zero, then \(\lim_{n \to \infty} \left| \frac{A_{n+1}}{A_n} \right| |x - c| = 0\) for every \(x\) and the ratio test tells us that the series \(\sum_{n=0}^{\infty} A_n(x - c)^n\) converges for every number \(x\). In this case we say that the series has an infinite radius of convergence.

\textit{By convention, when \((x - c)^0\) appears in a power series, it has value 1 for all values of \(x\), even \(x = c\).}
• If the $\left| \frac{A_{n+1}}{A_n} \right|$ tends to $+\infty$ as $n \to 0$, then $\lim_{n \to \infty} \left| \frac{A_{n+1}}{A_n} \right| |x - c| = +\infty$ for every $x \neq c$ and the ratio test tells us that the series $\sum_{n=0}^{\infty} A_n(x - c)^n$ diverges for every number $x \neq c$. As we have seen above, when $x = c$, the series reduces to $A_0 + 0 + 0 + 0 + 0 + \cdots$, which of course converges. In this case we say that the series has radius of convergence zero.

• If $\left| \frac{A_{n+1}}{A_n} \right|$ does not approach a limit as $n \to \infty$, then we learn nothing from the ratio test.

All of these possibilities do happen. We give an example of each below. But first, the concept of “radius of convergence” is important enough to warrant an official definition.

**Definition 3.5.2.**

(a) Let $0 < R < \infty$. If $\sum_{n=0}^{\infty} A_n(x - c)^n$ converges for $|x - c| < R$, and diverges for $|x - c| > R$, then we say that the series has radius of convergence $R$.

(b) If $\sum_{n=0}^{\infty} A_n(x - c)^n$ converges for every number $x$, we say that the series has an infinite radius of convergence.

(c) If $\sum_{n=0}^{\infty} A_n(x - c)^n$ diverges for every $x \neq c$, we say that the series has radius of convergence zero.

**Example 3.5.3**

We already know that, if $a \neq 0$, the geometric series $\sum_{n=0}^{\infty} ax^n$ converges when $|x| < 1$ and diverges when $|x| \geq 1$. So, in the terminology of Definition 3.5.2, the geometric series has radius of convergence $R = 1$. As a consistency check, we can also compute $R$ using (3.5.1).

The series $\sum_{n=0}^{\infty} ax^n$ has $A_n = a$. So

$$R = \left[ \lim_{n \to \infty} \left| \frac{A_{n+1}}{A_n} \right| \right]^{-1} = \left[ \lim_{n \to \infty} 1 \right]^{-1} = 1$$

as expected.

**Example 3.5.4**

Recall that $n! = 1 \times 2 \times 3 \times \cdots \times n$ is called “$n$ factorial”. By convention $0! = 1$.

The series $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ has $A_n = \frac{1}{n!}$. So

$$\lim_{n \to \infty} \left| \frac{A_{n+1}}{A_n} \right| = \lim_{n \to \infty} \frac{1/(n+1)!}{1/n!} = \lim_{n \to \infty} \frac{n!}{(n+1)!} = \lim_{n \to \infty} \frac{1 \times 2 \times 3 \times \cdots \times n \times (n+1)}{n!} = \lim_{n \to \infty} \frac{1}{n+1} = 0$$
and \( \sum_{n=0}^{\infty} \frac{x^n}{n!} \) has radius of convergence \( \infty \). It converges for every \( x \).

**Example 3.5.4**

The series \( \sum_{n=0}^{\infty} n!x^n \) has \( A_n = n! \). So

\[
\lim_{n \to \infty} \left| \frac{A_{n+1}}{A_n} \right| = \lim_{n \to \infty} \frac{(n+1)!}{n!} = \lim_{n \to \infty} \frac{1 \times 2 \times 3 \times 4 \times \cdots \times n \times (n+1)}{1 \times 2 \times 3 \times 4 \times \cdots \times n} = \lim_{n \to \infty} (n+1) = +\infty
\]

and \( \sum_{n=0}^{\infty} n!x^n \) has radius of convergence zero. It converges only for \( x = 0 \), where it takes the value \( 0! = 1 \).

**Example 3.5.5**

Let \( A_0 = 4 \) and, for each integer \( n \geq 1 \), let \( A_n \) be one plus the \( n^{th} \) decimal digit of \( \pi \). So every \( A_n \) is an integer between 1 and 10 and the series

\[
\sum_{n=0}^{\infty} A_n x^n = 4 + 2x + 5x^2 + 2x^3 + 6x^4 + 10x^5 + \cdots
\]

Because \( \pi \) is an irrational number \( \frac{A_{n+1}}{A_n} \) cannot have a limit as \( n \to \infty \). (If you don’t know why this is the case, don’t worry about it.) So the ratio test tells us nothing about the convergence of this series. But we can still figure out for which \( x \)'s it converges.

- Because every coefficient \( A_n \) is no bigger (in magnitude) than 10, the \( n^{th} \) term in our series obeys

\[
|A_n x^n| \leq 10|x|^n
\]

and so is smaller than the \( n^{th} \) term in the geometric series \( \sum_{n=0}^{\infty} 10|x|^n \). This geometric series converges if \( |x| < 1 \). So, by the comparison test, our series converges for \( |x| < 1 \) too.

- Since every \( A_n \) is at least one, the \( n^{th} \) term in our series obeys

\[
|A_n x^n| \geq |x|^n
\]

If \( |x| \geq 1 \), this \( a_n = A_n x^n \) cannot converge to zero as \( n \to \infty \), and our series diverges by the divergence test.

In conclusion, our series converges if and only if \( |x| < 1 \), and so has radius of convergence 1.

**Example 3.5.6**
Though we won’t prove it, it is true that every power series has a radius of convergence, whether or not the limit \( \lim_{n \to \infty} \left| \frac{A_{n+1}}{A_n} \right| \) exists.

**Theorem 3.5.7.**

Let \( \sum_{n=0}^{\infty} A_n (x - c)^n \) be a power series. Then one of the following alternatives must hold.

(i) The power series converges for every number \( x \). In this case we say that the radius of convergence is \( \infty \).

(ii) There is a number \( 0 < R < \infty \) such that the series converges for \( |x - c| < R \) and diverges for \( |x - c| > R \). Then \( R \) is called the radius of convergence.

(iii) The series converges for \( x = 0 \) and diverges for all \( x \neq 0 \). In this case, we say that the radius of convergence is \( 0 \).

**Definition 3.5.8.**

The set of \( x \)'s for which a power series converges is called the interval of convergence for the series.

Suppose that the power series \( \sum_{n=0}^{\infty} A_n (x - c)^n \) has radius of convergence \( R \). Then from Theorem 3.5.7, we have that

- if \( R = \infty \), then its interval of convergence is \( -\infty < x < \infty \), which is also denoted \( (-\infty, \infty) \), and
- if \( 0 < R < \infty \), then, because

\[
|x - c| < R \quad \text{if and only if} \quad -R < x - c < R
\]

\[
\text{and}\quad \text{if and only if} \quad c - R < x < c + R
\]

its interval of convergence must be one of

- \( c - R < x < c + R \), which is also denoted \( (c - R, c + R) \), or
- \( c - R \leq x < c + R \), which is also denoted \( [c - R, c + R) \), or
- \( c - R < x \leq c + R \), which is also denoted \( (c - R, c + R] \), or
- \( c - R \leq x \leq c + R \), which is also denoted \( [c - R, c + R] \),

and

- if \( R = 0 \), then its interval of convergence is just the point \( x = 0 \).
Note that knowing the radius of convergence, $R$ with $0 < R < \infty$, tells you nothing about whether or not the series converges when $|x - c| = R$, i.e. when $x = c \pm R$. The following example shows that all four possibilities occur.

Example 3.5.9

Let $p$ be any real number and consider the series $\sum_{n=0}^{\infty} \frac{x^n}{n^p}$. This series has $A_n = \frac{1}{n^p}$. Since

$$\lim_{n \to \infty} \left| \frac{A_{n+1}}{A_n} \right| = \lim_{n \to \infty} \frac{n^p}{(n+1)^p} = \lim_{n \to \infty} \frac{1}{(1 + 1/n)^p} = 1$$

the series has radius of convergence 1. So it certainly converges for $|x| < 1$ and diverges for $|x| > 1$. That just leaves $x = \pm 1$.

- When $x = 1$, the series reduces to $\sum_{n=0}^{\infty} \frac{1}{n^p}$. We know, from Example 3.3.5, that this series converges if and only if $p > 1$.

- When $x = -1$, the series reduces to $\sum_{n=0}^{\infty} \frac{(-1)^n}{n^p}$. By the alternating series test, Theorem 3.3.12, this series converges whenever $p > 0$ (so that $\frac{1}{n^p}$ tends to zero as $n$ tends to infinity). When $p \leq 0$ (so that $\frac{1}{n^p}$ does not tend to zero as $n$ tends to infinity), it diverges by the divergence test, Theorem 3.3.1.

So

(a) The power series $\sum_{n=0}^{\infty} x^n$ (i.e. $p = 0$) has interval of convergence $-1 < x < 1$.

(b) The power series $\sum_{n=0}^{\infty} \frac{x^n}{n}$ (i.e. $p = 1$) has interval of convergence $-1 \leq x < 1$.

(c) The power series $\sum_{n=0}^{\infty} \frac{(-1)^n}{n^p} x^n$ (i.e. $p = 1$) has interval of convergence $-1 < x \leq 1$.

(d) The power series $\sum_{n=0}^{\infty} \frac{x^n}{n^2}$ (i.e. $p = 2$) has interval of convergence $-1 \leq x \leq 1$.

Example 3.5.10

We are told that a certain power series with centre 3, that is, in powers of $(x - 3)$, converges at $x = 4$ and diverges at $x = 1$. The following figure provides a resume this given convergence data. Green dots mark the values of $x$ where the series is known to converge. (Recall that every power series converges at its centre.) The red dot marks the value of $x$ where the series is known to diverge.

Can we say more about the convergence and/or divergence of the series for other values of $x$? Yes! Certainly the series has some, at this stage still unknown, radius of convergence $R$ with

\[ 293 \]
the series converging at \( x \) if \( |x - 3| < R \) and
the series diverging at \( x \) if \( |x - 3| > R \).

We have also been told that

- the series converges when \( x = 4 \) so
  - \( x = 4 \) cannot obey \( |x - 3| > R \) so
  - \( x = 4 \) must obey \( |x - 3| \leq R \), i.e. \( |4 - 3| \leq R \), i.e. \( R \geq 1 \)
- the series diverges when \( x = 1 \) so
  - \( x = 1 \) cannot obey \( |x - 3| < R \) so
  - \( x = 1 \) must obey \( |x - 3| \geq R \), i.e. \( |1 - 3| \geq R \), i.e. \( R \leq 2 \)

We still don’t know \( R \) exactly. But we do know that \( 1 \leq R \leq 2 \). Consequently,

- since 1 is the smallest that \( R \) could be, the series certainly converges at \( x \) if \( |x - 3| < 1 \), i.e. if \( 2 < x < 4 \) and
- since 2 is the largest that \( R \) could be, the series certainly diverges at \( x \) if \( |x - 3| > 2 \), i.e. if \( x > 5 \) or if \( x < 1 \).

The following figure provides a resume of all of this convergence data — there is convergence at green \( x \)’s and divergence at red \( x \)’s.

### 3.5.2 Working With Power Series

Here is a theorem that can be used help build power series representations for complicated functions from power series representations of simple functions.
**Theorem 3.5.11** (Operations on Power Series). Assume that the functions $f(x)$ and $g(x)$ are given by the power series

$$f(x) = \sum_{n=0}^{\infty} A_n (x-c)^n \quad g(x) = \sum_{n=0}^{\infty} B_n (x-c)^n$$

for all $x$ obeying $|x-c| < R$. In particular, we are assuming that both power series have radius of convergence at least $R$. Also let $K$ be a nonzero constant. Then

$$f(x) + g(x) = \sum_{n=0}^{\infty} [A_n + B_n] (x-c)^n$$

$$Kf(x) = \sum_{n=0}^{\infty} K A_n (x-c)^n$$

$$(x-c)^N f(x) = \sum_{n=0}^{\infty} A_n (x-c)^{n+N} \quad \text{for any integer } N \geq 1$$

$$= \sum_{k=N}^{\infty} A_{k-N} (x-c)^k \quad \text{where } k = n + N$$

$$f'(x) = \sum_{n=0}^{\infty} n A_n (x-c)^{n-1}$$

$$\int_c^x f(t) \, dt = \sum_{n=0}^{\infty} A_n \frac{(x-c)^{n+1}}{n+1}$$

$$\int f(x) \, dx = \left[ \sum_{n=0}^{\infty} A_n \frac{(x-c)^{n+1}}{n+1} \right] + C \quad \text{with } C \text{ an arbitrary constant}$$

for all $x$ obeying $|x-c| < R$. In particular, the radius of convergence of each of the six power series on the right hand sides is at least $R$. In fact, if $R$ is the radius of convergence of $\sum_{n=0}^{\infty} A_n (x-c)^n$, then $R$ is also the radius of convergence of all of the above right hand sides, with the possible exception of $\sum_{n=0}^{\infty} [A_n + B_n] (x-c)^n$.

We’ll now use this theorem to build power series representations for a bunch of functions out of the one simple power series representation that we know — the geometric series

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \quad \text{for all } |x| < 1$$

**Example 3.5.12** ($\frac{1}{1-x^2}$)

Find a power series representation for $\frac{1}{1-x^2}$. 

295
Solution. The secret to finding power series representations for a good many functions is to manipulate them into a form in which \( \frac{1}{1 - y} \) appears and use the geometric series representation \( \frac{1}{1 - y} = \sum_{n=0}^{\infty} y^n \). We have deliberately renamed the variable to \( y \) here — it does not have to be \( x \). We can use that strategy to find a power series expansion for \( \frac{1}{1 - x^2} \) — we just have to recognize that \( \frac{1}{1 - x^2} \) is the same as \( \frac{1}{1 - y} \) if we set \( y \) to \( x^2 \).

\[
\frac{1}{1 - x^2} = \left. \frac{1}{1 - y} \right|_{y=x^2} = \left[ \sum_{n=0}^{\infty} y^n \right]_{y=x^2} = \sum_{n=0}^{\infty} (x^2)^n = \sum_{n=0}^{\infty} x^{2n} = 1 + x^2 + x^4 + x^6 + \cdots
\]

This is a perfectly good power series. There is nothing wrong with the power of \( x \) being \( 2n \). In fact, you should try to always write power series in forms that are as easy to understand as possible. The geometric series that we used at the end of the first line converges for

\[
|y| < 1 \iff |x|^2 < 1 \iff |x| < 1
\]

So our power series has radius of convergence 1 and interval of convergence \(-1 < x < 1\).

Example 3.5.12

Find a power series representation for \( \frac{x}{2 + x^2} \).

Solution. This example is just a more algebraically involved variant of the last one. Again, the strategy is to manipulate \( \frac{x}{2 + x^2} \) into a form in which \( \frac{1}{1 - y} \) appears.

\[
x \left( \frac{1}{2 + x^2} \right) = \frac{x}{2} \left( \frac{1}{1 + (-x^2/2)} \right) = \frac{x}{2} \left. \frac{1}{1 - (-x^2/2)} \right|_{y=-x^2/2} = \frac{x}{2} \left[ \sum_{n=0}^{\infty} (-1)^n y^n \right]_{y=-x^2/2} = \frac{x}{2} \sum_{n=0}^{\infty} (-1)^n \frac{(-x^2/2)^n}{2n+1}
\]

The geometric series that we used in the second line converges for

\[
|y| < 1 \iff |-x^2/2| < 1 \iff |x|^2 < 2 \iff |x| < \sqrt{2}
\]
So our power series has radius of convergence $\sqrt{2}$ and interval of convergence $-\sqrt{2} < x < \sqrt{2}$.

Example 3.5.13

Find a power series representation for $\frac{1}{(1-x)^2}$.

Solution. Once again the trick is to express $\frac{1}{(1-x)^2}$ in terms of $\frac{1}{1-x}$.

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

$$= \frac{d}{dx} \sum_{n=0}^{\infty} x^n$$

$$= \sum_{n=1}^{\infty} nx^{n-1} \quad \text{by Theorem 3.5.11}$$

Note that the $n = 0$ term has disappeared because, for $n = 0$

$$\frac{d}{dx} x^n = \frac{d}{dx} 1 = 0$$

Again, Theorem 3.5.11 guarantees that the radius of convergence is exactly one (the radius of convergence of the geometric series $\sum_{n=0}^{\infty} x^n$). By the divergence test, the series diverges for $x = \pm 1$. So our power series has radius of convergence 1 and interval of convergence $-1 < x < 1$.

Example 3.5.14

Find a power series representation for $\log(1+x)$.

Solution. Recall that $\frac{d}{dx} \log(1+x) = \frac{1}{1+x}$ so that $\log(1+t)$ is an antiderivative of $\frac{1}{1+t}$ and

$$\log(1+x) = \int_{0}^{x} \frac{dt}{1+t} = \int_{0}^{x} \left[ \sum_{n=0}^{\infty} (-t)^n \right] dt$$

$$= \sum_{n=0}^{\infty} \int_{0}^{x} (-t)^n dt \quad \text{by Theorem 3.5.11}$$

$$= \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1}$$

$$= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$$

Example 3.5.15
Theorem 3.5.11 guarantees that the radius of convergence is exactly one (the radius of convergence of the geometric series \( \sum_{n=0}^{\infty} (-t)^n \)). When \( x = -1 \) our series reduces to minus \( \sum_{n=0}^{\infty} \frac{1}{n+1} \), which is the harmonic series and so diverges. That’s no surprise — \( \log(1 + (-1)) = \log 0 = -\infty \). When \( x = 1 \), the series converges by the alternating series test. It is possible to prove, though we won’t do so here, that the sum is \( \log 2 \). So the interval of convergence is \( -1 < x \leq 1 \).

Example 3.5.15

Example 3.5.16 (arctan \( x \))

Find a power series representation for \( \arctan x \).

Solution. Recall that \( \frac{d}{dx} \arctan x = \frac{1}{1+x^2} \) so that \( \arctan t \) is an antiderivative of \( \frac{1}{1+t^2} \) and

\[
\arctan x = \int_0^x \frac{dt}{1+t^2} = \int_0^x \left[ \sum_{n=0}^{\infty} (-t^2)^n \right] dt = \sum_{n=0}^{\infty} \int_0^x (-1)^n t^{2n} dt \\
= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} \\
= x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots
\]

Theorem 3.5.11 guarantees that the radius of convergence is exactly one (the radius of convergence of the geometric series \( \sum_{n=0}^{\infty} (-t^2)^n \)). When \( x = \pm 1 \), the series converges by the alternating series test. So the interval of convergence is \( -1 \leq x \leq 1 \). It is possible to prove, though once again we won’t do so here, that when \( x = \pm 1 \), the series converges to \( \arctan(\pm 1) = \pm \frac{\pi}{4} \).

3.6 Taylor Series

3.6.1 Extending Taylor Polynomials

Recall that Taylor polynomials provide a hierarchy of approximations to a given function \( f(x) \) near a given point \( a \).

- The crudest approximation is the constant approximation \( f(x) \approx f(a) \).
- Then comes the linear, or tangent line, approximation \( f(x) \approx f(a) + f'(a) (x - a) \).
- Then comes the quadratic approximation \( f(x) \approx f(a) + f'(a) (x - a) + \frac{1}{2} f''(a) (x - a)^2 \).
- In general, the Taylor polynomial of degree \( n \), for the function \( f(x) \) about the expansion point \( a \), is the polynomial, \( T_n(x) \), determined by the requirements that
Applying (3.6.1) with \( f \) is the Taylor polynomial of degree \( n \). That is, \( f \) and \( T_n \) have the same derivatives at \( a \), up to order \( n \). Explicitly,

\[
f(x) \approx T_n(x) = f(a) + f'(a) (x - a) + \frac{1}{2} f''(a) (x - a)^2 + \cdots + \frac{1}{n!} f^{(n)}(a) (x - a)^n
\]

These are of course approximations — often very good approximations near \( x = a \) — but still just approximations. Can we get exact representations by taking the limit as \( n \to \infty \)? That’s the question we’ll consider now.

Fix a real number \( a \) and suppose that all derivatives of the function \( f(x) \) exist. Then, for any natural number \( n \),

\[
f(x) = T_n(x) + E_n(x) \tag{3.6.1}
\]

where

\[
T_n(x) = f(a) + f'(a) (x - a) + \cdots + \frac{1}{n!} f^{(n)}(a) (x - a)^n \tag{3.6.1a}
\]

is the Taylor polynomial of degree \( n \) for the function \( f(x) \) and expansion point \( a \), and \( E_n(x) = f(x) - T_n(x) \) is the error introduced when we approximate \( f(x) \) by the polynomial \( T_n(x) \). It is true, though we won’t prove it, that

\[
E_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(c) (x - a)^{n+1} \tag{3.6.1b}
\]

for some (usually unknown) \( c \) strictly between \( a \) and \( x \).

If it happens that, for some \( x \), \( E_n(x) \) tends to zero as \( n \to \infty \), then, for that \( x \), we have the exact formula

\[
f(x) = \lim_{n \to \infty} T_n(x)
\]

for \( f(x) \). This is usually written

\[
f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(a) (x - a)^n \tag{3.6.2}
\]

and is called the Taylor series of \( f(x) \) with expansion point \( a \). (When \( a = 0 \) it is also called the Maclaurin series of \( f(x) \).) It is a power series representation for \( f(x) \).

**Example 3.6.1 (Exponential Series)**

This happens with the exponential function \( f(x) = e^x \). We’ll first find \( f^{(m)}(0) \) for all integers \( m \geq 0 \).

\[
\begin{align*}
f(x) &= e^x & \Rightarrow & & f'(x) &= e^x & \Rightarrow & & f''(x) &= e^x & \cdots \\
\quad f(0) &= e^0 = 1 & \Rightarrow & & f'(0) &= e^0 = 1 & \Rightarrow & & f''(0) &= e^0 = 1 & \cdots 
\end{align*}
\]

Applying (3.6.1) with \( f(x) = e^x \) and \( a = 0 \), and using that \( f^{(m)}(a) = e^a = e^0 = 1 \) for all \( m \),

\[
e^x = f(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + E_n(x) \tag{3.6.3}
\]
We shall see, in Example 3.6.2 below, that, for any fixed \( x \), \( \lim_{n \to \infty} E_n(x) = 0 \). Consequently, for all \( x \),
\[
e^x = \lim_{n \to \infty} \left[ 1 + x + \frac{1}{2} x^2 + \frac{1}{3!} x^3 + \cdots + \frac{1}{n!} x^n \right] = \sum_{n=0}^\infty \frac{1}{n!} x^n \quad (3.6.4)
\]

**Example 3.6.1**

We have already seen, in Example 3.6.1, that
\[
e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + E_n(x) \quad (3.6.5)
\]

By (3.6.1b)
\[
E_n(x) = \frac{1}{(n+1)!} e^c x^{n+1}
\]

for some (unknown) \( c \) between 0 and \( x \). Fix any real number \( x \). We’ll now show that \( E_n(x) \) converges to zero as \( n \to \infty \).

As \( c \) runs from 0 to \( x \), \( e^c \) runs from \( e^0 = 1 \) to \( e^x \). In particular, \( e^c \) is always between 1 and \( e^x \) and so is no bigger the larger of the two numbers 1 and \( e^x \), which is turn is smaller than \( 1 + e^x \). Thus the error term
\[
|E_n(x)| = \left| \frac{e^c x^{n+1}}{(n+1)!} \right| \leq \left| e^x + 1 \right| \frac{|x|^{n+1}}{(n+1)!}
\]

Let’s call \( e_n(x) = \frac{|x|^{n+1}}{(n+1)!} \). We claim that as \( n \) increases towards infinity, \( e_n(x) \) decreases (quickly) towards zero. To see this, let’s compare \( e_n(x) \) and \( e_{n+1}(x) \).
\[
\frac{e_{n+1}(x)}{e_n(x)} = \frac{|x|^{n+2}}{(n+2)!} \frac{(n+1)!}{|x|^{n+1}} = \frac{|x|}{n+2}
\]

So, when \( n \) is bigger than, for example \( 2|x| \), we have \( \frac{e_{n+1}(x)}{e_n(x)} < \frac{1}{2} \). That is, increasing the index on \( e_n(x) \) by one decreases the size of \( e_n(x) \) by a factor of at least two. As a result \( e_n(x) \) must tend to zero as \( n \to \infty \). Consequently, for all \( x \), \( \lim_{n \to \infty} E_n(x) = 0 \), as claimed, and we really have
\[
e^x = \lim_{n \to \infty} \left[ 1 + x + \frac{1}{2} x^2 + \frac{1}{3!} x^3 + \cdots + \frac{1}{n!} x^n \right] = \sum_{n=0}^\infty \frac{1}{n!} x^n
\]

**Example 3.6.2**

(Optional — Why \( \sum_{n=0}^\infty \frac{1}{n!} x^n \) is \( e^x \))

In this example, we will again show that the \( E_n(x) \) of (3.6.5) converges to zero as \( n \) tends
to infinity. But we shall do so using a tricky method in which $E_n(x)$ never explicitly appears. We know, from Example 3.5.4, that the series $\sum_{n=0}^{\infty} \frac{1}{n!} x^n$ converges for all values of $x$ to some function $f(x)$. We wish to verify that $f(x)$ really is $e^x$. To do so, define $f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} x^n$ and make the following two observations.

- By Theorem 3.5.11,
  \[
  \frac{df}{dx} = \frac{d}{dx} \sum_{n=0}^{\infty} \frac{1}{n!} x^n = \sum_{n=1}^{\infty} \frac{n}{n!} x^{n-1} = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} x^{n-1} = f(x)
  \]

- $f(0) = 1$

That is, $f(x)$ is the solution of the initial value problem
\[
\frac{dy}{dx} = y \quad y(0) = 1
\]

We know from Theorem 2.4.3, with $a = 1, b = 0$ and $y(0) = 1$, that this forces $f(x) = e^x$.

Example 3.6.3

Example 3.6.4 (Sine and Cosine Series)

The trigonometric functions $\sin x$ and $\cos x$ also have widely used Taylor series expansions about $a = 0$. To find them, we first, compute all derivatives at general $x$.

\[
\begin{align*}
  f(x) &= \sin x & f'(x) &= \cos x & f''(x) &= -\sin x & f^{(3)}(x) &= -\cos x & f^{(4)}(x) &= \sin x & \cdots \\
  f(x) &= \cos x & f'(x) &= -\sin x & f''(x) &= -\cos x & f^{(3)}(x) &= \sin x & f^{(4)}(x) &= \cos x & \cdots
\end{align*}
\]

The pattern starts over again with the fourth derivative being the same as the original function. Now set $x = a = 0$.

\[
\begin{align*}
  f(0) &= \sin 0 = 0 & f'(0) &= \cos 0 = 1 & f''(0) &= 0 & f^{(3)}(0) &= -1 & f^{(4)}(0) &= 0 & \cdots \\
  f(0) &= \cos 0 = 1 & f'(0) &= 0 & f''(0) &= 0 & f^{(3)}(0) &= 0 & f^{(4)}(0) &= 1 & \cdots
\end{align*}
\]

(3.6.7)

For $\sin x$, all even numbered derivatives (at $x = 0$) are zero. The odd numbered derivatives alternate between 1 and $-1$. For $\cos x$, all odd numbered derivatives are zero. The even numbered derivatives alternate between 1 and $-1$. So, the Taylor polynomials that best approximate $\sin x$ and $\cos x$ near $x = a = 0$ are

\[
\begin{align*}
  \sin x &\approx x - \frac{1}{3!} x^3 + \frac{1}{5!} x^5 - \cdots \\
  \cos x &\approx 1 - \frac{1}{2!} x^2 + \frac{1}{4!} x^4 - \cdots
\end{align*}
\]
We shall see, in Example 3.6.5 below, that, for both $f(x) = \sin x$ and $f(x) = \cos x$, we have $\lim_{n \to \infty} E_n(x) = 0$ so that

$$f(x) = \lim_{n \to \infty} \left[ f(0) + f'(0) x + \cdots + \frac{1}{n!} f^{(n)}(0) x^n \right]$$

Reviewing (3.6.7), we conclude that, for all $x$,

$$\sin x = x - \frac{1}{3!} x^3 + \frac{1}{5!} x^5 - \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \tag{3.6.8}$$

$$\cos x = 1 - \frac{1}{2!} x^2 + \frac{1}{4!} x^4 - \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$$

Let $f(x)$ be either $\sin x$ or $\cos x$. Reviewing (3.6.6) we see that every derivative of $\sin x$ and $\cos x$ is one of $\pm \sin x$ and $\pm \cos x$. Consequently, when we apply (3.6.1b) we always have $|f^{(n+1)}(c)| \leq 1$ and hence $|E_n(x)| \leq \frac{|x|^{n+1}}{(n+1)!}$. We have already seen, in Example 3.6.1, that $\frac{|x|^{n+1}}{(n+1)!}$ (which we called $e_n(x)$ in Example 3.6.1) converges to zero as $n \to \infty$. Consequently, for both $f(x) = \sin x$ and $f(x) = \cos x$, we have $\lim_{n \to \infty} E_n(x) = 0$ and

$$f(x) = \lim_{n \to \infty} \left[ f(0) + f'(0) x + \cdots + \frac{1}{n!} f^{(n)}(0) x^n \right]$$

We have developed power series representations for a number of important functions. Here is a theorem that summarizes them.
### Theorem 3.6.6

\[
e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots \quad \text{for all } -\infty < x < \infty
\]

\[
\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots \quad \text{for all } -\infty < x < \infty
\]

\[
\cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots \quad \text{for all } -\infty < x < \infty
\]

\[
\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \cdots \quad \text{for all } -1 < x < 1
\]

\[
\log(1+x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} - \cdots \quad \text{for all } -1 < x \leq 1
\]

\[
arctan x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots \quad \text{for all } -1 \leq x \leq 1
\]

### Example 3.6.7 (Computing the number \(\pi\))

There are numerous methods for computing \(\pi\) to any desired degree of accuracy. Many of them use the Taylor expansion\(^\text{17}\)

\[
arctan x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} \quad (3.6.9)
\]

that we derived in Example 3.5.16. One of the simplest uses \(\tan \frac{\pi}{6} = \frac{1}{\sqrt{3}}\):

\[
\pi = 6 \arctan \left( \frac{1}{\sqrt{3}} \right) = 6 \sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1} \frac{1}{(\sqrt{3})^{2n+1}}
\]

\[
= 2\sqrt{3} \sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1} \frac{1}{3^n}
\]

\[
= 2\sqrt{3} \left( 1 - \frac{1}{3 \times 3} + \frac{1}{5 \times 9} - \frac{1}{7 \times 27} + \frac{1}{9 \times 81} - \frac{1}{11 \times 243} + \cdots \right)
\]

This is an alternating series. So, by Theorem 3.3.12, the error we introduce by truncating it is between zero and the first term dropped. For example, if we keep ten terms, stopping at \(n = 9\), we get \(\pi = 3.141591\) (to 6 decimal places) with an error between zero and

\[
\frac{2\sqrt{3}}{21 \times 3^{10}} < 3 \times 10^{-6}
\]

\(^{17}\) This is indeed the Taylor expansion, with centre 0, (i.e. Maclaurin expansion) of \(\arctan x\), even though we did not use (3.6.1) to derive it.
This is just one of very many ways to compute $\pi$. Another one, which still uses (3.6.9) but is much more efficient, is

$$\pi = 16 \arctan \frac{1}{5} - 4 \arctan \frac{1}{239}$$

This formula was used by John Machin in 1706 to compute $\pi$ to 100 decimal digits.

**Example 3.6.7**

The *error function*

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \, dt$$

is used in computing “bell curve” probabilities. The indefinite integral of the integrand $e^{-t^2}$ cannot be expressed in terms of standard functions. But we can still evaluate the integral to within any desired degree of accuracy by using the Taylor expansion of the exponential.

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \left[ \sum_{n=0}^{\infty} \frac{(-t^2)^n}{n!} \right] \, dt$$

$$= \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)n!}$$

For example, for the bell curve, the probability of being within one standard deviation of the mean (if you don’t know what this means, ignore the words) is

$$\text{erf} \left( \frac{1}{\sqrt{2}} \right) = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} (-1)^n \frac{\left( \frac{1}{\sqrt{2}} \right)^{2n+1}}{(2n+1)n!} = \frac{2}{\sqrt{2\pi}} \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n+1)2^n n!}$$

$$= \sqrt{\frac{2}{\pi}} \left( 1 - \frac{1}{3 \times 2} + \frac{1}{5 \times 2^2 \times 2} - \frac{1}{7 \times 2^3 \times 3!} + \frac{1}{9 \times 2^4 \times 4!} - \cdots \right)$$

This is yet another alternating series. If we keep five terms, stopping at $n = 4$, we get 0.68271 (to 5 decimal places) with, by Theorem 3.3.12 again, an error between zero and the first dropped term, which is minus

$$\sqrt{\frac{2}{\pi}} \frac{1}{11 \times 2^5 \times 5!} < 2 \times 10^{-5}$$

**Example 3.6.8**

Evaluate

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n}$$
Solution. There are not very many series that can be easily evaluated exactly. But occasionally one encounters a series that can be evaluated simply by realizing that it is exactly one of the series in Theorem 3.6.6, just with a specific value of $x$. That is the case for the given series, which is

$$
\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots
$$

Comparing this with the series given in Theorem 3.6.6, we see that

$$
\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots
$$

$$
= \left[ x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} - \cdots \right]_{x=1}
$$

$$
= \log(1 + x)_{x=1}
$$

$$
= \log 2
$$

Example 3.6.10 (Optional — Computing the number $e$)

We have seen, in (3.6.3), that

$$
e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \frac{1}{(n+1)!}e^c x^{n+1}
$$

(3.6.10)

for some $c$ between 0 and $x$. We can use this to find approximate values for the number $e$, with any desired degree of accuracy. Just setting $x = 1$ in (3.6.10) gives

$$
e = 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{n!} + \frac{1}{(n+1)!}e^c
$$

(3.6.11)

for some $c$ between 0 and 1. Since $e^c$ increases as $c$ increases, this says that $1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{n!}$ is an approximate value for $e$ with error at most $\frac{e}{(n+1)!}$. The only problem with this error bound is that it contains the number $e$, which we do not know. Fortunately, we can again use (3.6.11) to get a simple upper bound on how big $e$ can be. Just setting $n = 2$ in (3.6.11), and again using that $e^c \leq e$, gives

$$
e \leq 1 + 1 + \frac{1}{2!} + \frac{e}{3!} \implies (1 - \frac{1}{3!})e = (1 - \frac{1}{6})e \leq 1 + 1 + \frac{1}{2!} = \frac{5}{2} \implies e \leq \frac{5}{2} \times \frac{6}{5} = 3
$$

So we now know that $1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{n!}$ is an approximate value for $e$ with error at most $\frac{3}{(n+1)!}$. For example, when $n = 9$, $\frac{3}{10!} < 10^{-6}$ so that

$$
1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{9!} \leq e \leq 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{9!} + 10^{-6}
$$

with

$$
1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \frac{1}{6!} + \frac{1}{7!} + \frac{1}{8!} + \frac{1}{9!} = 1 + 1 + 0.5 + 0.16 + 0.0416 + 0.0083 + 0.00138 + 0.0001984 + 0.0000248 + 0.0000028 = 2.718282
$$
to six decimal places.

Example 3.6.10

Let \( f(x) = \sin(2x^3) \). Find \( f^{(15)}(0) \), the fifteenth derivative of \( f \) at \( x = 0 \).

**Solution.** This is a bit of a trick question. We could of course use the product and chain rules to directly apply fifteen derivatives and then set \( x = 0 \), but that would be extremely tedious. There is a much more efficient approach that exploits two pieces of knowledge that we have.

- From (3.6.2), we see that the coefficient of \((x - a)^n\) in the Taylor series of \( f(x) \) with expansion point \( a \) is exactly \( \frac{1}{n!} f^{(n)}(a) \). So \( f^{(n)}(a) \) is exactly \( n! \) times the coefficient of \((x - a)^n\) in the Taylor series of \( f(x) \) with expansion point \( a \).
- We know, or at least can easily find, the Taylor series for \( \sin(2x^3) \).

Let’s apply that strategy.

- First, we know that

\[
\sin x = x - \frac{1}{3!} x^3 + \frac{1}{5!} x^5 - \cdots
\]

- Just replacing \( x \) by \( 2x^3 \), we have

\[
\sin(2x^3) = 2x^3 - \frac{1}{3!} (2x^3)^3 + \frac{1}{5!} (2x^3)^5 - \cdots
\]

\[
= 2x^3 - \frac{8}{3!} x^9 + \frac{2^5}{5!} x^{15} - \cdots
\]

- So the coefficient of \( x^{15} \) in the Taylor series of \( f(x) = \sin(2x^3) \) with expansion point \( a = 0 \) is \( \frac{2^5}{5!} \)

and we have

\[
f^{(15)}(0) = 15! \times \frac{2^5}{5!} = 348,713,164,800
\]

Example 3.6.11

3.6.2 Evaluating Limits using Taylor Expansions

Taylor polynomials provide a good way to understand the behaviour of a function near a specified point and so are useful for evaluating complicated limits. Here are some examples.

Example 3.6.12

In this example, we’ll start with a relatively simple limit, namely

\[
\lim_{{x \to 0}} \frac{\sin x}{x}
\]
The first thing to notice about this limit is that, as \( x \) tends to zero, both the numerator, \( \sin x \), and the denominator, \( x \), tend to 0. So we may not evaluate the limit of the ratio by simply dividing the limits of the numerator and denominator. To find the limit, or show that it does not exist, we are going to have to exhibit a cancellation between the numerator and the denominator. Let’s start by taking a closer look at the numerator. By Example 3.6.4,

\[
\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots
\]

Consequently

\[
\frac{\sin x}{x} = 1 - \frac{1}{3!}x^2 + \frac{1}{5!}x^4 - \cdots
\]

Every term in this series, except for the very first term, is proportional to a strictly positive power of \( x \). Consequently, as \( x \) tends to zero, all terms in this series, except for the very first term, tend to zero. In fact the sum of all terms, starting with the second term, also tends to zero. That is,

\[
\lim_{x \to 0} \left[ -\frac{1}{3!}x^2 + \frac{1}{5!}x^4 - \cdots \right] = 0
\]

We won’t justify that statement here, but it will be justified in the following (optional) subsection. So

\[
\lim_{x \to 0} \frac{\sin x}{x} = \lim_{x \to 0} \left[ 1 - \frac{1}{3!}x^2 + \frac{1}{5!}x^4 - \cdots \right]
= 1 + \lim_{x \to 0} \left[ -\frac{1}{3!}x^2 + \frac{1}{5!}x^4 - \cdots \right]
= 1
\]

The limit in the previous example can also be evaluated relatively easily using l’Hôpital’s rule, if you know it. While the following limit can also, in principal, be evaluated using l’Hôpital’s rule, it is much more efficient to use Taylor series.

In this example we evaluate

\[
\lim_{x \to 0} \frac{\arctan x - x}{\sin x - x}
\]

Once again, the first thing to notice about this limit is that, as \( x \) tends to zero, the numerator tends to \( \arctan 0 - 0 \), which is 0, and the denominator tends to \( \sin 0 - 0 \), which is also 0. So we may not evaluate the limit of the ratio by simply dividing the limits of the numerator and denominator. Again, to find the limit, or show that it does not exist, we are going to have to exhibit a cancellation between the numerator and the denominator. To get a more detailed understanding of the behaviour of the numerator and denominator near \( x = 0 \), we find their Taylor expansions. By Example 3.5.16,

\[
\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots
\]
so the numerator
\[
\arctan x - x = -\frac{x^3}{3} + \frac{x^5}{5} - \cdots
\]

By Example 3.6.4,
\[
\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots
\]
so the denominator
\[
\sin x - x = -\frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots
\]
and the ratio
\[
\frac{\arctan x - x}{\sin x - x} = \frac{-\frac{x^3}{3} + \frac{x^5}{5} - \cdots}{-\frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots}
\]
Notice that every term in both the numerator and the denominator contains a common factor of \( x^3 \), which we can cancel out.

\[
\frac{\arctan x - x}{\sin x - x} = \frac{-\frac{1}{3} + \frac{x^2}{5} - \cdots}{-\frac{1}{3!} + \frac{1}{5!}x^2 - \cdots}
\]

As \( x \) tends to zero,
- the numerator tends to \(-\frac{1}{3}\), which is not 0, and
- the denominator tends to \(-\frac{1}{3!} = -\frac{1}{6}\), which is also not 0.

so we may now legitimately evaluate the limit of the ratio by simply dividing the limits of the numerator and denominator.

\[
\lim_{x \to 0} \frac{\arctan x - x}{\sin x - x} = \frac{\lim_{x \to 0} -\frac{1}{3} + \frac{x^2}{5} - \cdots}{\lim_{x \to 0} -\frac{1}{3!} + \frac{1}{5!}x^2 - \cdots}
\]
\[
= \frac{\lim_{x \to 0} \left[ -\frac{1}{3} + \frac{x^2}{5} - \cdots \right]}{\lim_{x \to 0} \left[ -\frac{1}{3!} + \frac{1}{5!}x^2 - \cdots \right]}
\]
\[
= \frac{-\frac{1}{3}}{-\frac{1}{3!}}
\]
\[
= 2
\]

Example 3.6.13

3.6.3 Optional — The Big O Notation

In Example 3.6.12 we used, without justification, that, as \( x \) tends to zero, not only does every term in
\[
\frac{\sin x}{x} - 1 = -\frac{1}{3!}x^2 + \frac{1}{5!}x^4 - \cdots = \sum_{n=1}^{\infty} (-1)^n \frac{1}{(2n+1)!} x^{2n}
\]
converge to zero, but in fact the sum of all infinitely many terms also converges to zero. We did something similar twice in Example 3.6.13, once in computing the limit of the numerator and once in computing the limit of the denominator.

We’ll now develop some machinery that provides the justification. We start by recalling, from (3.6.1), that if, for some natural number \( n \), the function \( f(x) \) has \( n + 1 \) derivatives near the point \( a \), then

\[
f(x) = T_n(x) + E_n(x)
\]

where

\[
T_n(x) = f(a) + f'(a) (x - a) + \cdots + \frac{1}{n!} f^{(n)}(a) (x - a)^n
\]

is the Taylor polynomial of degree \( n \) for the function \( f(x) \) and expansion point \( a \) and

\[
E_n(x) = f(x) - T_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(c) (x - a)^{n+1}
\]

is the error introduced when we approximate \( f(x) \) by the polynomial \( T_n(x) \). Here \( c \) is some unknown number between \( a \) and \( x \). As \( c \) is not known, we do not know exactly what the error \( E_n(x) \) is. But that is usually not a problem. In taking the limit \( x \to a \), we are only interested in \( x \)'s that are very close to \( a \), and when \( x \) is very close, \( c \) must also be very close to \( a \). As long as \( f^{(n+1)}(x) \) is continuous at \( a \), \( f^{(n+1)}(c) \) must approach \( f^{(n)}(a) \) as \( x \to a \). In particular there must be constants \( M, D > 0 \) such that \( |f^{(n+1)}(c)| \leq M \) for all \( c \)'s within a distance \( D \) of \( a \). If so, there is another constant \( C \) (namely \( \frac{M}{(n+1)!} \)) such that

\[
|E_n(x)| \leq C|x - a|^{n+1} \quad \text{whenever } |x - a| \leq D
\]

There is some notation for this behaviour.

**Definition 3.6.14 (Big O).**

Let \( a \) and \( m \) be real numbers. We say “\( F(x) \) is of order \( |x - a|^m \) near \( a \)” and we write \( F(x) = O(|x - a|^m) \) if there exist constants \( C, D > 0 \) such that

\[
|F(x)| \leq C|x - a|^m \quad \text{whenever } |x - a| \leq D \tag{3.6.12}
\]

Whenever \( O(|x - a|^m) \) appears in an algebraic expression, it just stands for some (unknown) function \( \bar{F}(x) \) that obeys (3.6.12). This is called “big O” notation.

**Example 3.6.15**

Let \( f(x) = \sin x \) and \( a = 0 \). Then

\[
\begin{align*}
f(x) &= \sin x & f'(x) &= \cos x & f''(x) &= - \sin x & f^{(3)}(x) &= - \cos x & f^{(4)}(x) &= \sin x & \cdots \\
f(0) &= 0 & f'(0) &= 1 & f''(0) &= 0 & f^{(3)}(0) &= -1 & f^{(4)}(0) &= 0 & \cdots
\end{align*}
\]

and the pattern repeats. Thus \( |f^{(n+1)}(c)| \leq 1 \) for all real numbers \( c \) and all natural numbers \( n \). So the Taylor polynomial of, for example, degree 3 and its error term are

\[
\sin x = x - \frac{1}{3!} x^3 + \frac{\cos c}{5!} x^5 \\
= x - \frac{1}{3!} x^3 + O(|x|^5)
\]
under Definition 3.6.14, with $C = \frac{1}{3!}$ and any $D > 0$. Similarly, for any natural number $n$,

\[
\begin{align*}
\sin x &= x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots + (-1)^n \frac{1}{(2n+1)!} x^{2n+1} + O(|x|^{2n+3}) \\
\cos x &= 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots + (-1)^n \frac{1}{(2n)!} x^{2n} + O(|x|^{2n+2})
\end{align*}
\] (3.6.13) (3.6.14)

Example 3.6.15

Let $n$ be any natural number. Since $\frac{d^m}{dx^m} e^x = e^x$ for every integer $m \geq 0$,

\[
e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \frac{e^c}{(n+1)!} x^{n+1}
\]

for some $c$ between 0 and $x$. If, for example, $|x| \leq 1$, then $|e^c| \leq e$, so that the error term

\[
\left| \frac{e^c}{(n+1)!} x^{n+1} \right| \leq C|x|^{n+1}
\]

with $C = \frac{e}{(n+1)!}$ whenever $|x| \leq 1$

So, under Definition 3.6.14, with $C = \frac{e}{(n+1)!}$ and $D = 1$,

\[
e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + O(|x|^{n+1})
\] (3.6.15)

Example 3.6.16

Let $f(x) = \log(1 + x)$ and $a = 0$. Then

\[
\begin{align*}
f'(x) &= \frac{1}{1+x} \\
f''(x) &= -\frac{1}{(1+x)^2} \\
f^{(3)}(x) &= -\frac{2}{(1+x)^3} \\
f^{(4)}(x) &= -\frac{6}{(1+x)^4} \\
f^{(5)}(x) &= -\frac{24}{(1+x)^5}
\end{align*}
\]

\[
f'(0) = 1 \\
f''(0) = -1 \\
f^{(3)}(0) = 2 \\
f^{(4)}(0) = -3 \\
f^{(5)}(0) = 4!
\]

We can see a pattern for $f^{(n)}(x)$ forming here — $f^{(n)}(x)$ is a sign times a ratio with

- the sign being $+$ when $n$ is odd and being $-$ when $n$ is even. So the sign is $(-1)^{n-1}$.
- The denominator is a power of $(1 + x)$. The power is just $n$.
- The numerator is a product $2 \times 3 \times 4 \times \cdots$. The last integer in the power is $n - 1$, at least for $n \geq 2$. So the product, for $n \geq 2$, is $2 \times 3 \times 4 \times \cdots \times (n - 1)$. The notation $n!$, read “$n$ factorial”, means $1 \times 2 \times 3 \times \cdots \times n$, so the numerator is $(n - 1)!$, at least for $n \geq 2$. By convention, $0! = 1$, so the numerator is $(n - 1)!$ for $n = 1$ too.

Thus, for any natural number $n$,

\[
f^{(n)}(x) = (-1)^{n-1} \frac{(n-1)!}{(1+x)^n} f^{(n)}(0) x^n = (-1)^{n-1} \frac{(n-1)!}{n!} x^n = (-1)^{n-1} \frac{x^n}{n}
\]

so

\[
\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n-1} \frac{x^n}{n} + E_n(x)
\]
with
\[ E_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(c) (x-a)^{n+1} = (-1)^n \frac{1}{(n+1)(1+c)^{n+1}} x^{n+1} \]

If we choose, for example \( D = \frac{1}{2} \), then for any \( x \) obeying \( |x| \leq D = \frac{1}{2} \), we have \( |c| \leq \frac{1}{2} \) and \( |1+c| \geq \frac{1}{2} \) so that
\[ |E_n(x)| \leq \frac{1}{(n+1)(1/2)^{n+1}} |x|^{n+1} = O(|x|^{n+1}) \]
under Definition 3.6.14, with \( C = \frac{2^{n+1}}{n+1} \) and \( D = \frac{1}{2} \). Thus we may write
\[
\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n-1} \frac{x^n}{n} + O(|x|^{n+1}) \tag{3.6.16}
\]

**Example 3.6.17**

Remark 3.6.18. The big O notation has a few properties that are useful in computations and taking limits. All follow immediately from Definition 3.6.14.

1. If \( p > 0 \), then
\[
\lim_{x \to 0} O(|x|^p) = 0
\]

2. For any real numbers \( p \) and \( q \),
\[
O(|x|^p) O(|x|^q) = O(|x|^{p+q})
\]
(This is just because \( C|x|^p \times C'|x|^q = (CC')|x|^{p+q} \). In particular,
\[
ax^m O(|x|^p) = O(|x|^{p+m})
\]
for any constant \( a \) and any integer \( m \).

3. For any real numbers \( p \) and \( q \),
\[
O(|x|^p) + O(|x|^q) = O(|x|^\min\{p,q\})
\]
(For example, if \( p = 2 \) and \( q = 5 \), then \( C|x|^2 + C'|x|^5 = (C + C'|x|^3)|x|^2 \leq (C + C')|x|^2 \) whenever \( |x| \leq 1 \).)

4. For any real numbers \( p \) and \( q \) with \( p > q \), any function which is \( O(|x|^p) \) is also \( O(|x|^q) \) because \( C|x|^p = C|x|^{p-q}|x|^q \leq C|x|^q \) whenever \( |x| \leq 1 \).

### 3.6.4 Optional — Evaluating Limits Using Taylor Expansions — More Examples

**Example 3.6.19** (Example 3.6.12 revisited)

In this example, we’ll return to the limit
\[
\lim_{x \to 0} \frac{\sin x}{x}
\]
of Example 3.6.12 and treat it more carefully. By Example 3.6.15,

\[ \sin x = x - \frac{1}{3!} x^3 + O(|x|^5) \]

That is, for small \( x \), \( \sin x \) is the same as \( x - \frac{1}{3!} x^3 \), up to an error that is bounded by some constant times \( |x|^5 \). So, dividing by \( x \), \( \frac{\sin x}{x} \) is the same as \( 1 - \frac{1}{3!} x^2 \), up to an error that is bounded by some constant times \( x^4 \). (See Remark 3.6.18.b.) That is

\[ \frac{\sin x}{x} = 1 - \frac{1}{3!} x^2 + O(x^4) \]

But any function that is bounded by some constant times \( x^4 \) (for all \( x \) smaller than some constant \( D > 0 \)) necessarily tends to 0 as \( x \to 0 \). Thus

\[ \lim_{x \to 0} \frac{\sin x}{x} = \lim_{x \to 0} \left[ 1 - \frac{1}{3!} x^2 + O(x^4) \right] = \lim_{x \to 0} \left[ 1 - \frac{1}{3!} x^2 \right] = 1 \]

Reviewing the above computation, we see that we did a little more work than we had to. It wasn’t necessary to keep track of the \(-\frac{1}{3!} x^3\) contribution to \( \sin x \) so carefully. We could have just said that

\[ \sin x = x + O(|x|^3) \]

so that

\[ \lim_{x \to 0} \frac{\sin x}{x} = \lim_{x \to 0} \frac{x + O(|x|^3)}{x} = \lim_{x \to 0} \left[ 1 + O(x^2) \right] = 1 \]

We’ll spend a little time in the later, more complicated, examples learning how to choose the number of terms we keep in our Taylor expansions so as to make our computations as efficient as possible.

Example 3.6.19

Example 3.6.20

In this example, we’ll use the Taylor polynomial of Example 3.6.17 to evaluate \( \lim_{x \to 0} \frac{\log(1+x)}{x} \)

and \( \lim_{x \to 0} \frac{\log(1+x)}{x} \). The Taylor expansion (3.6.16) with \( n = 1 \) tells us that

\[ \log(1+x) = x + O(|x|^2) \]

That is, for small \( x \), \( \log(1+x) \) is the same as \( x \), up to an error that is bounded by some constant times \( x^2 \). So, dividing by \( x \), \( \frac{1}{x} \log(1+x) \) is the same as 1, up to an error that is bounded by some constant times \( |x| \). That is

\[ \frac{1}{x} \log(1+x) = 1 + O(|x|) \]

But any function that is bounded by some constant times \( |x| \), for all \( x \) smaller than some constant \( D > 0 \), necessarily tends to 0 as \( x \to 0 \). Thus

\[ \lim_{x \to 0} \frac{\log(1+x)}{x} = \lim_{x \to 0} \frac{x + O(|x|^2)}{x} = \lim_{x \to 0} \left[ 1 + O(|x|) \right] = 1 \]
and
\[ \lim_{x \to 0} (1 + x)^{a/x} = \lim_{x \to 0} e^{a/x \log(1 + x)} = \lim_{x \to 0} e^{a/x [x + O(|x|^2)]} = \lim_{x \to 0} e^{a + O(|x|)} = e^a. \]

Here we have used that if \( F(x) = O(|x|^2) \), that is if \( |F(x)| \leq C|x|^2 \) for some constant \( C \), then \( \frac{d}{dx} F(x) \leq C' |x| \) for the new constant \( C' = |a|C \), so that \( F(x) = O(|x|) \). We have also used that the exponential is continuous — as \( x \) tends to zero, the exponent of \( e^{a + O(|x|)} \) tends to \( a \) so that \( e^{a + O(|x|)} \) tends to \( e^a \).

Example 3.6.20

In this example, we’ll evaluate the harder limit
\[ \lim_{x \to 0} \frac{\cos x - 1 + \frac{1}{2} x \sin x}{[\log(1 + x)]^4} \]

The first thing to notice about this limit is that, as \( x \) tends to zero, the numerator, which is \( \cos x - 1 + \frac{1}{2} x \sin x \), tends to \( \cos 0 - 1 + \frac{1}{2} \cdot 0 \cdot \sin 0 = 0 \) and the denominator \( [\log(1 + x)]^4 \) tends to \( [\log(1 + 0)]^4 = 0 \) too. So both the numerator and denominator tend to zero and we may not simply evaluate the limit of the ratio by taking the limits of the numerator and denominator and dividing. To find the limit, or show that it does not exist, we are going to have to exhibit a cancellation between the numerator and the denominator. To develop a strategy for evaluating this limit, let’s do a “little scratch work”, starting by taking a closer look at the denominator. By Example 3.6.17,
\[ \log(1 + x) = x + O(x^2) \]

This tells us that \( \log(1 + x) \) looks a lot like \( x \) for very small \( x \). So the denominator \( [x + O(x^2)]^4 \) looks a lot like \( x^4 \) for very small \( x \). Now, what about the numerator?

- If the numerator looks like some constant times \( x^p \) with \( p > 4 \), for very small \( x \), then the ratio will look like the constant times \( \frac{x^p}{x^4} = x^{p-4} \) and will tend to 0 as \( x \) tends to zero.
- If the numerator looks like some constant times \( x^p \) with \( p < 4 \), for very small \( x \), then the ratio will look like the constant times \( \frac{x^p}{x^4} = x^{p-4} \) and will tend to infinity, and in particular diverge, as \( x \) tends to zero.
- If the numerator looks like \( Cx^4 \), for very small \( x \), then the ratio will look like \( \frac{C}{x^4} = C \) and will tend to \( C \) as \( x \) tends to zero.

The moral of the above “scratch work” is that we need to know the behaviour of the numerator, for small \( x \), up to order \( x^4 \). Any contributions of order \( x^p \) with \( p > 4 \) may be put into error terms \( O(|x|^p) \). Now we are ready to evaluate the limit. Using Examples 3.6.15 and 3.6.17,
\[
\lim_{x \to 0} \frac{\cos x - 1 + \frac{1}{2} x \sin x}{[\log(1 + x)]^4} = \lim_{x \to 0} \frac{[1 - \frac{1}{2} x^2 + \frac{1}{4!} x^4 + O(x^6)] - 1 + \frac{1}{2} x [x - \frac{1}{3} x^3 + O(|x|^5)]}{[x + O(x^2)]^4} \]
\[
= \lim_{x \to 0} \frac{(\frac{1}{4!} - \frac{1}{2 \times 3!}) x^4 + O(x^6) + \frac{5}{2} O(|x|^5)}{[x + O(x^2)]^4} \]

313
\[
\begin{align*}
&= \lim_{x \to 0} \frac{(\frac{1}{4} - \frac{1}{2 \cdot 4}) x^4 + O(x^6) + O(x^6)}{[x + O(x^2)]^4} \quad \text{by Remark 3.6.18, part 2.} \\
&= \lim_{x \to 0} \frac{(\frac{1}{4} - \frac{1}{2 \cdot 4}) x^4 + O(x^6)}{[x + x O(|x|)]^4} \quad \text{by Remark 3.6.18, parts 2, 3.} \\
&= \lim_{x \to 0} \frac{(\frac{1}{4} - \frac{1}{2 \cdot 4}) x^4 + x^4 O(x^2)}{x^4[1 + O(|x|)]^4} \quad \text{by Remark 3.6.18, part 2.} \\
&= \lim_{x \to 0} \frac{(\frac{1}{4} - \frac{1}{2 \cdot 4}) + O(x^2)}{[1 + O(|x|)]^4} \quad \text{by Remark 3.6.18 part 1.} \\
&= \frac{1}{4!} - \frac{1}{2 \times 3!} \\
&= \frac{1}{3!} \left( \frac{1}{4} - \frac{1}{2} \right) = -\frac{1}{4!}.
\end{align*}
\]

**Example 3.6.21**

In this example we’ll evaluate another harder limit, namely

\[
\lim_{x \to 0} \frac{\log \left( \frac{\sin x}{x} \right)}{x^2}
\]

The first thing to notice about this limit is that, as \(x\) tends to zero, the denominator \(x^2\) tends to 0. So, yet again, to find the limit, we are going to have to show that the numerator also tends to 0 and we are going to have to exhibit a cancellation between the numerator and the denominator.

Because the denominator is \(x^2\) any terms in the numerator, \(\log \left( \frac{\sin x}{x} \right)\) that are of order \(x^3\) or higher will contribute terms in the ratio \(\frac{\log \left( \frac{\sin x}{x} \right)}{x^2}\) that are of order \(x\) or higher. Those terms in the ratio will converge to zero as \(x \to 0\). The moral of this discussion is that we need to compute \(\log \left( \frac{\sin x}{x} \right)\) to order \(x^2\) with errors of order \(x^3\). Now we saw, in Example 3.6.19, that

\[
\frac{\sin x}{x} = 1 - \frac{1}{3!} x^2 + O(x^4)
\]

We also saw, in (3.6.16) with \(n = 1\), that

\[
\log(1 + X) = X + O(X^2) \quad (3.6.17)
\]

Substituting \(X = -\frac{1}{3!} x^2 + O(x^4)\), and using that \(X^2 = O(x^4)\) (by Remark 3.6.18, parts 2 and 3), we have that the numerator

\[
\log \left( \frac{\sin x}{x} \right) = \log(1 + X) = X + O(X^2) = -\frac{1}{3!} x^2 + O(x^4)
\]

and the limit

\[
\lim_{x \to 0} \frac{\log \left( \frac{\sin x}{x} \right)}{x^2} = \lim_{x \to 0} \frac{-\frac{1}{3!} x^2 + O(x^4)}{x^2} = \lim_{x \to 0} \left[ -\frac{1}{3!} + O(x^2) \right] = -\frac{1}{3!} = -\frac{1}{6}
\]
Evaluate
\[
\lim_{x \to 0} \frac{e^{x^2} - \cos x}{\log(1 + x) - \sin x}
\]

**Solution.**

*Step 1:* Find the limit of the denominator.
\[
\lim_{x \to 0} [\log(1 + x) - \sin x] = \log(1 + 0) - \sin 0 = 0
\]
This tells us that we can’t evaluate the limit just by finding the limits of the numerator and denominator separately and then dividing.

*Step 2:* Determine the leading order behaviour of the denominator near \(x = 0\). By (3.6.16) and (3.6.13),
\[
\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots
\]
\[
\sin x = x - \frac{1}{6!} x^3 + \frac{1}{3!} x^5 - \cdots
\]
Subtracting, the denominator
\[
\log(1 + x) - \sin x = -\frac{x^2}{2} + \left(\frac{1}{3} + \frac{1}{3!}\right)x^3 + \cdots
\]
This tells us that, for \(x\) near zero, the denominator is \(-\frac{x^2}{2}\) (that’s the leading order term) plus contributions that are of order \(x^3\) and smaller. That is
\[
\log(1 + x) - \sin x = -\frac{x^2}{2} + O(|x|^3)
\]

*Step 3:* Determine the behaviour of the numerator near \(x = 0\) to order \(x^2\) with errors of order \(x^3\) and smaller (just like the denominator). By (3.6.15)
\[
e^x = 1 + X + O(X^2)
\]
Substituting \(X = x^2\)
\[
e^{x^2} = 1 + x^2 + O(x^4)
\]
\[
\cos x = 1 - \frac{x^2}{2} + O(x^4)
\]
by (3.6.14). Subtracting, the numerator
\[
e^{x^2} - \cos x = \frac{3}{2} x^2 + O(x^4)
\]

*Step 4:* Evaluate the limit.
\[
\lim_{x \to 0} \frac{e^{x^2} - \cos x}{\log(1 + x) - \sin x} = \lim_{x \to 0} \frac{\frac{3}{2} x^2 + O(x^4)}{-\frac{x^2}{2} + O(|x|^3)} = \lim_{x \to 0} \frac{\frac{3}{2} + O(x^2)}{-\frac{1}{2} + O(|x|)} = \frac{3/2}{-1/2} = -3
\]
3.7 Optional — Rational and irrational numbers

In this optional section we shall use series techniques to look a little at rationality and irrationality of real numbers. We shall see the following results.

- A real number is rational (i.e. a ratio of two integers) if and only if its decimal expansion is eventually periodic. "Eventually periodic" means that, if we denote the \(n\)th decimal place by \(d_n\), then there are two positive integers \(k\) and \(p\) such that \(d_{n+p} = d_n\) whenever \(n > k\). So the part of the decimal expansion after the decimal point looks like

\[
\ldots a_1a_2a_3\ldots a_k b_1b_2\ldots b_p b_1b_2\ldots b_p b_1b_2\ldots b_p \ldots
\]

It is possible that a finite number of decimal places right after the decimal point do not participate in the periodicity. It is also possible that \(p = 1\) and \(b_1 = 0\), so that the decimal expansion ends with an infinite string of zeros.

- \(e\) is irrational.
- \(\pi\) is irrational.

Decimal expansions of rational numbers

We start by showing that a real number is rational if and only if its decimal expansion is eventually periodic. We need only consider the expansions of numbers \(0 < x < 1\). If a number is negative then we can just multiply it by \(-1\) and not change the expansion. Similarly if the number is larger than 1 then we can just subtract off the integer part of the number and leave the expansion unchanged.

- Eventually periodic implies rational

Let us assume that a number \(0 < x < 1\) has a decimal expansion that is eventually periodic. Hence we can write

\[
x = 0.\underbrace{a_1a_2a_3\ldots a_k}_{\text{a has at most } k \text{ digits}} b_1b_2\ldots b_p \underbrace{b_1b_2\ldots b_p}_{\text{b has at most } p \text{ digits}} \ldots
\]

Let \(\alpha = a_1a_2a_3\ldots a_k\) and \(\beta = b_1b_2\ldots b_p\). In particular, \(\alpha\) has at most \(k\) digits and \(\beta\) has at most \(p\) digits. Then we can (carefully) write

\[
x = \frac{\alpha}{10^k} + \frac{\beta}{10^{k+p}} + \frac{\beta}{10^{k+2p}} + \frac{\beta}{10^{k+3p}} + \ldots
\]

This sum is just a geometric series (see Example 3.2.4) and we can evaluate it:

\[
= \frac{\alpha}{10^k} + \frac{\beta}{10^{k+p}} \sum_{j=0}^{\infty} 10^{-p} = \frac{\alpha}{10^k} + \frac{\beta}{10^{k+p}} \cdot \frac{1}{1 - 10^{-p}} = \frac{\alpha}{10^k} + \frac{\beta}{10^{k+p}} + \frac{1}{10^p - 1}
\]

This is a ratio of integers, so \(x\) is a rational number.
Rational implies eventually periodic

Let $0 < x < 1$ be rational with $x = \frac{a}{b}$, where $a$ and $b$ are positive integers. We wish to show that $x$’s decimal expansion is eventually periodic. Start by looking at the last formula we derived in the “eventually periodic implies rational” subsection. If we can express the denominator $b$ in the form $\frac{10^k(10^p - 1)}{q}$ with $k$, $p$ and $q$ integers, we will be in business because $\frac{a}{b} = \frac{aq}{10^k(10^p - 1)}$. From this we can generate the desired decimal expansion by running the argument of the last subsection backwards. So we want to find integers $k$, $p$, $q$ such that $10^k + p - 10^k = b \cdot q$. To do so consider the powers of 10 up to $10^b$:

$$1, 10^1, 10^2, 10^3, \ldots, 10^b$$

For each $j = 0, 1, 2, \ldots, b$, find integers $c_j$ and $0 \leq r_j < b$ so that

$$10^j = b \cdot c_j + r_j$$

To do so, start with $10^j$ and repeatedly subtract $b$ from it until the remainder drops strictly below $b$. The $r_j$’s can take at most $b + 1$ different values, namely $0, 1, 2, \ldots, b - 1$, and we now have $b + 1$ $r_j$’s, namely $r_0, r_1, \ldots, r_b$. So we must be able to find two powers of 10 which give the same remainder. That is there must be $0 \leq k < l \leq b$ so that $r_k = r_l$. Hence

$$10^l - 10^k = (bc_l + r_l) - (bc_k + r_k)$$

$$= b(c_l - c_k)$$

and we have

$$b = \frac{10^k(10^p - 1)}{q}$$

where $p = l - k$ and $q = c_l - c_k$ are both strictly positive integers, since $l > k$ so that $10^l - 10^k > 0$. Thus we can write

$$\frac{a}{b} = \frac{aq}{10^k(10^p - 1)}$$

Next divide the numerator $aq$ by $10^p - 1$ and compute the remainder. That is, write $aq = a(10^p - 1) + \beta$ with $0 \leq \beta < 10^p - 1$. Notice that $0 \leq a < 10^k$, as otherwise $x = \frac{a}{b} \geq 1$. That is, $a$ has at most $k$ digits and $\beta$ has at most $p$ digits. This, finally, gives us

$$x = \frac{a}{b} = \frac{a(10^p - 1) + \beta}{10^k(10^p - 1)}$$

$$= \frac{\alpha}{10^k} + \frac{\beta}{10^k(10^p - 1)}$$

$$= \frac{\alpha}{10^k} + \frac{\beta}{10^k + p(1 - 10^{-p})}$$

$$= \frac{\alpha}{10^k} + \frac{\beta}{10^k + p} \sum_{j=0}^{\infty} 10^{-pj}$$

which gives the required eventually periodic expansion.

---

18 This is an application of the pigeon hole principle — the very simple but surprisingly useful idea that if you have $n$ items which you have to put in $m$ boxes, and if $n > m$, then at least one box must contain more than one item.
Irrationality of $e$

We will give 2 proofs that the number $e$ is irrational, the first due to Fourier (1768–1830) and the second due to Pennisi (1918–2010). Both are proofs by contradiction — we first assume that $e$ is rational and then show that this implies a contradiction. In both cases we reach the contradiction by showing that a given quantity (related to the series expression for $e$) must be both a positive integer and also strictly less than 1.

Proof 1

This proof is due to Fourier. Let us assume that the number $e$ is rational so we can write it as

$$e = \frac{a}{b}$$

where $a, b$ are positive integers. Using the Maclaurin series for $e^x$ we have

$$\frac{a}{b} = e^1 = \sum_{n=0}^{\infty} \frac{1}{n!}$$

Now multiply both sides by $b!$ to get

$$\frac{a}{b} \cdot b! = \sum_{n=0}^{\infty} \frac{b!}{n!}$$

The left-hand side of this expression is an integer. We complete the proof by showing that the right-hand side cannot be an integer (and hence that we have a contradiction).

First split the series on the right-hand side into two piece as follows

$$\sum_{n=0}^{\infty} \frac{b!}{n!} = \sum_{n=0}^{b} \frac{b!}{n!} + \sum_{n=b+1}^{\infty} \frac{b!}{n!} = A + B$$

The first sum, $A$, is finite sum of integers:

$$A = \sum_{n=0}^{b} \frac{b!}{n!} = \sum_{n=0}^{b} (n+1)(n+2) \cdots (b-1)b.$$

Consequently $A$ must be an integer. Notice that we simplified the ratio of factorials using the fact that when $b \geq n$ we have

$$\frac{b!}{n!} = \frac{1 \cdot 2 \cdots n \cdot (n+1)(n+2) \cdots (b-1)b}{1 \cdot 2 \cdots n} = (n+1)(n+2) \cdots (b-1)b.$$

Proof by contradiction is a standard and very powerful method of proof in mathematics. It relies on the law of the excluded middle which states that any given mathematical statement $P$ is either true or false. Because of this, if we can show that the statement $P$ being false implies something contradictory — like $1 = 0$ or $a > a$ — then we can conclude that $P$ must be true. The interested reader can certainly find many examples (and a far more detailed explanation) using their favourite search engine.
Now we turn to the second sum. Since it is a sum of strictly positive terms we must have

\[ B > 0 \]

We complete the proof by showing that \( B < 1 \). To do this we bound each term from above:

\[
\frac{b!}{n!} = \frac{1}{(b+1)(b+2) \cdots (n-1)n}
\]

\[
\leq \frac{1}{(b+1)(b+1) \cdots (b+1)(b+1)} = \frac{1}{(b+1)^{n-b}}
\]

Indeed the inequality is strict except when \( n = b + 1 \). Hence we have that

\[ B < \sum_{n=b+1}^{\infty} \frac{1}{(b+1)^{n-b}} = \frac{1}{b+1} + \frac{1}{(b+1)^2} + \frac{1}{(b+1)^3} + \cdots \]

This is just a geometric series (see Example 3.2.4) and equals

\[
= \frac{1}{b+1} \left( 1 - \frac{1}{b+1} \right) = \frac{1}{b+1 - 1} = \frac{1}{b}
\]

And since \( b \) is a positive integer, we have shown that

\[ 0 < B < 1 \]

and thus \( B \) cannot be an integer.

Thus we have that

\[
\underbrace{\frac{a}{b}}_{\text{integer}} + \underbrace{B}_{\text{not integer}} = \underbrace{A}_{\text{integer}}
\]

which gives a contradiction. Thus \( e \) cannot be rational.

Proof 2

This proof is due to Pennisi (1953). Let us (again) assume that the number \( e \) is rational. Hence it can be written as

\[ e = \frac{a}{b'} \]
where \( a, b \) are positive integers. This means that we can write
\[
e^{-1} = \frac{b}{a},
\]
Using the Maclaurin series for \( e^x \) we have
\[
\frac{b}{a} = e^{-1} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!}
\]
Before we do anything else, we multiply both sides by \((-1)^{a+1}a\) — this might seem a little strange at this point, but the reason will become clear as we proceed through the proof. The expression is now
\[
(-1)^{a+1} \frac{a!}{a} = \sum_{n=0}^{\infty} \frac{(-1)^{n+a+1}a!}{n!}
\]
The left-hand side of the expression is an integer. We again complete the proof by showing that the right-hand side cannot be an integer.

We split the series on the right-hand side into two pieces:
\[
= \sum_{n=0}^{a} \frac{(-1)^{n+a+1}a!}{n!} + \sum_{n=a+1}^{\infty} \frac{(-1)^{n+a+1}a!}{n!}
\]
We will show that \( A \) is an integer while \( 0 < B < 1 \); this gives the required contradiction.

Every term in the sum \( A \) is an integer. To see this we simplify the ratio of factorials as we did in the previous proof:
\[
A = \sum_{n=0}^{a} \frac{(-1)^{n+a+1}a!}{n!} = \frac{(-1)^{n+a+1}(n+1)(n+2)\ldots(a-1)a}{a!}
\]
Let us now examine the series \( B \). Again clean up the ratio of factorials:
\[
B = \sum_{n=a+1}^{\infty} \frac{(-1)^{n+a+1}a!}{n!} = \sum_{n=a+1}^{\infty} \frac{(-1)^{n+a+1}}{(a+1)(a+2)\ldots(n-1)n}
\]
Hence \( B \) is an alternating series of decreasing terms and by the alternating series test (Theorem 3.3.12) it converges. Further, it must converge to a number between its first and second partial sums (see the discussion before Theorem 3.3.12). Hence the right-hand side lies between
\[
\frac{1}{a+1} \quad \text{and} \quad \frac{1}{a+1} - \frac{1}{(a+1)(a+2)} = \frac{1}{a+2}
\]
Since \( a \) is a positive integer the above tells us that \( B \) converges to a real number strictly greater than 0 and strictly less than 1. Hence it cannot be an integer.

This gives us a contradiction and hence \( e \) cannot be rational.
**Irrationality of** $\pi$

This proof is due to Niven (1946) and doesn’t require any mathematics beyond the level of this course. Much like the proofs above we will start by assuming that $\pi$ is rational and then reach a contradiction. Again this contradiction will be that a given quantity must be an integer but at the same time must lie strictly between 0 and 1.

Assume that $\pi$ is a rational number and so can be written as $\pi = \frac{a}{b}$ with $a, b$ positive integers. Now let $n$ be a positive integer and define the polynomial

$$f(x) = \frac{x^n(a - bx)^n}{n!}.$$ 

It is certainly not immediately obvious why and how Niven chose this polynomial, but you will see that it has been very carefully crafted to make the proof work. In particular we will show — under our assumption that $\pi$ is rational — that, if $n$ is really big, then

$$I_n = \int_0^\pi f(x) \sin(x) dx$$

is an integer and it also lies strictly between 0 and 1, giving the required contradiction.

**Bounding the integral**

Consider again the polynomial

$$f(x) = \frac{x^n(a - bx)^n}{n!}.$$ 

Notice that

$$f(0) = 0$$

$$f(\pi) = f(a/b) = 0.$$ 

Furthermore, for $0 \leq x \leq \pi = a/b$, we have $x \leq a/b$ and $a - bx \leq a$ so that

$$0 \leq x(a - bx) \leq a^2/b.$$ 

We could work out a more precise upper bound, but this one is sufficient for the analysis that follows. Hence

$$0 \leq f(x) \leq \left(\frac{a^2}{b}\right)^n \frac{1}{n!}$$ 

We also know that for $0 \leq x \leq \pi = a/b$, $0 \leq \sin(x) \leq 1$. Thus

$$0 \leq f(x) \sin(x) \leq \left(\frac{a^2}{b}\right)^n \frac{1}{n!}$$

for all $0 \leq x \leq 1$. Using this inequality we bound

$$0 < I_n = \int_0^\pi f(x) \sin(x) dx < \left(\frac{a^2}{b}\right)^n \frac{1}{n!}.$$ 

We will later show that, if $n$ is really big, then $\left(\frac{a^2}{b}\right)^n \frac{1}{n!} < 1$. We’ll first show, starting now, that $I_n$ is an integer.

---

You got lots of practice finding the maximum and minimum values of continuous functions on closed intervals when you took calculus last term.
Integration by parts

In order to show that the value of this integral is an integer we will use integration by parts. You have already practiced using integration by parts to integrate quantities like

\[ \int x^2 \sin(x) \, dx \]

and this integral isn’t much different. For the moment let us just use the fact that \( f(x) \) is a polynomial of degree \( 2n \). Using integration by parts with \( u = f(x) \), \( dv = \sin(x) \) and \( v = -\cos(x) \) gives us

\[ \int f(x) \sin(x) \, dx = -f(x) \cos(x) + \int f'(x) \cos(x) \, dx \]

Use integration by parts again with \( u = f'(x) \), \( dv = \cos(x) \) and \( v = \sin(x) \).

\[ = -f(x) \cos(x) + f'(x) \sin(x) - \int f''(x) \sin(x) \, dx \]

Use integration by parts yet again, with \( u = f''(x) \), \( dv = \sin(x) \) and \( v = -\cos(x) \).

\[ = -f(x) \cos(x) + f'(x) \sin(x) + f''(x) \cos(x) - \int f'''(x) \cos(x) \, dx \]

And now we can see the pattern; we get alternating signs, and then derivatives multiplied by sines and cosines:

\[ \int f(x) \sin(x) \, dx = \cos(x) \left( -f(x) + f''(x) - f^{(4)}(x) + f^{(6)}(x) - \cdots \right) \]
\[ + \sin(x) \left( f'(x) - f'''(x) + f^{(5)}(x) - f^{(7)}(x) + \cdots \right) \]

This terminates at the \( 2n \)th derivative since \( f(x) \) is a polynomial of degree \( 2n \). We can check this computation by differentiating the terms on the right-hand side:

\[
\frac{d}{dx} \left( \cos(x) \left( -f(x) + f''(x) - f^{(4)}(x) + f^{(6)}(x) - \cdots \right) \right)
\]
\[= -\sin(x) \left( -f(x) + f''(x) - f^{(4)}(x) + f^{(6)}(x) - \cdots \right) \]
\[+ \cos(x) \left( -f'(x) + f'''(x) - f^{(5)}(x) + f^{(7)}(x) - \cdots \right) \]

and similarly

\[
\frac{d}{dx} \left( \sin(x) \left( f'(x) - f'''(x) + f^{(5)}(x) - f^{(7)}(x) + \cdots \right) \right)
\]
\[= \cos(x) \left( f'(x) - f'''(x) + f^{(5)}(x) - f^{(7)}(x) + \cdots \right) \]
\[+ \sin(x) \left( f''(x) - f^{(4)}(x) + f^{(6)}(x) - \cdots \right) \]
When we add these two expressions together all the terms cancel except \( f(x) \sin(x) \), as required.

Now when we take the definite integral from 0 to \( \pi \), all the sine terms give 0 because
\[
\sin(0) = \sin(\pi) = 0.
\]
Since \( \cos(\pi) = -1 \) and \( \cos(0) = 1 \), we are just left with:
\[
\int_0^\pi f(x) \sin(x) \, dx = \left( f(0) - f''(0) + f^{(4)}(0) - f^{(6)}(0) + \cdots + (-1)^n f^{(2n)}(0) \right)
+ \left( f(\pi) - f''(\pi) + f^{(4)}(\pi) - f^{(6)}(\pi) + \cdots + (-1)^n f^{(2n)}(\pi) \right)
\]
So to show that \( I_n \) is an integer, it now suffices to show that \( f^{(j)}(0) \) and \( f^{(j)}(\pi) \) are integers.

\textbf{The derivatives are integers}

Recall that
\[
f(x) = \frac{x^n(a-bx)^n}{n!}
\]
and expand it:
\[
f(x) = \frac{c_0}{n!} x^0 + \frac{c_1}{n!} x^1 + \cdots + \frac{c_n}{n!} x^n + \cdots + \frac{c_{2n}}{n!} x^{2n}
\]
All the \( c_j \) are integers, and clearly \( c_j = 0 \) for all \( j = 0, 1, \ldots, n-1 \), because of the factor \( x^n \) in \( f(x) \).

Now take the \( k^{th} \) derivative and set \( x = 0 \). Note that, if \( j < k \), then \( \frac{d^k}{dx^k} x^j = 0 \) for all \( x \) and, if \( j > k \), then \( \frac{d^k}{dx^k} x^j \) is some number times \( x^{j-k} \) which evaluates to zero when we set \( x = 0 \). So
\[
f^{(k)}(0) = \frac{d^k}{dx^k} \left( \frac{c_k x^k}{k!} \right) = k! \frac{c_k}{n!}
\]
If \( k < n \), then this is zero since \( c_k = 0 \). If \( k > n \), this is an integer because \( c_k \) is an integer and \( k! / n! = (n+1)(n+2) \cdots (k-1)k \) is an integer. If \( k = n \), then \( f^{(k)}(0) = c_n \) is again an integer. Thus all the derivatives of \( f(x) \) evaluated at \( x = 0 \) are integers.

But what about the derivatives at \( \pi = a/b \)? To see this, we can make use of a handy symmetry. Notice that
\[
f(x) = f(\pi - x) = f(a/b - x)
\]
You can confirm this by just grinding through the algebra:
\[
f(x) = \frac{x^n(a-bx)^n}{n!} \quad \text{now replace } x \text{ with } a/b - x
\]
\[
f(a/b - x) = \frac{(a/b - x)^n(a - b(a/b - x))^n}{n!} \quad \text{start cleaning this up:}
\]
\[
= \frac{(a-bx)^n(a-a + bx)^n}{n!}
\]
\[
= \frac{(a-bx)^n(bx)^n}{n!}
\]
\[
= \frac{(a-bx)^n x^n}{n!} = f(x)
\]
Using this symmetry (and the chain rule) we see that

\[ f'(x) = -f'(\pi - x) \]

and if we keep differentiating

\[ f^{(k)}(x) = (-1)^k f^{(k)}(\pi - x) \]

Setting \( x = 0 \) in this tells us that

\[ f^{(k)}(0) = (-1)^k f^{(k)}(\pi) \]

So because all the derivatives at \( x = 0 \) are integers, we know that all the derivatives at \( x = \pi \) are also integers.

Hence the integral we are interested in

\[
\int_0^\pi f(x) \sin(x) \, dx
\]

must be an integer.

Putting it together

Based on our assumption that \( \pi = a/b \) is rational, we have shown that the integral

\[
I_n = \int_0^\pi \frac{x^n(a - bx)}{n!} \sin(x) \, dx
\]

satisfies

\[
0 < I_n < \left(\frac{a^2}{b}\right)^n \frac{1}{n!}
\]

and also that \( I_n \) is an integer.

We are, however, free to choose \( n \) to be any positive integer we want. If we take \( n \) to be very large — in particular much much larger than \( a \) — then \( n! \) will be much much larger than \( a^{2n} \) (we showed this in Example 3.6.2), and consequently

\[
0 < I_n < \left(\frac{a^2}{b}\right)^n \frac{1}{n!} < 1
\]

Which means that the integral cannot be an integer. This gives the required contradiction, showing that \( \pi \) is irrational.
This chapter is really split into three parts.

- Sections A.1 to A.11 contains results that we expect you to understand and know.
- Then Section A.14 contains results that we don’t expect you to memorise, but that we think you should be able to quickly derive from other results you know.
- The remaining sections contain some material (that may be new to you) that is related to topics covered in the main body of these notes.

### A.1 Similar triangles

Two triangles $T_1, T_2$ are similar when

- (AAA — angle angle angle) The angles of $T_1$ are the same as the angles of $T_2$.

- (SSS — side side side) The ratios of the side lengths are the same. That is

$$\frac{A}{a} = \frac{B}{b} = \frac{C}{c}$$

- (SAS — side angle side) Two sides have lengths in the same ratio and the angle between them is the same. For example

$$\frac{A}{a} = \frac{C}{c} \text{ and angle } \beta \text{ is same}$$
A.2 Pythagoras

For a right-angled triangle the length of the hypotenuse is related to the lengths of the other two sides by

\[(\text{adjacent})^2 + (\text{opposite})^2 = (\text{hypotenuse})^2\]

A.3 Trigonometry — definitions

\[
\begin{align*}
\sin \theta &= \frac{\text{opposite}}{\text{hypotenuse}} \\
\cos \theta &= \frac{\text{adjacent}}{\text{hypotenuse}} \\
\tan \theta &= \frac{\text{opposite}}{\text{adjacent}} \\
csc \theta &= \frac{1}{\sin \theta} \\
sec \theta &= \frac{1}{\cos \theta} \\
cot \theta &= \frac{1}{\tan \theta}
\end{align*}
\]

A.4 Radians, arcs and sectors

For a circle of radius \(r\) and angle of \(\theta\) radians:

- Arc length \(L(\theta) = r\theta\).
- Area of sector \(A(\theta) = \frac{\theta}{2}r^2\).
A.5 Trigonometry — graphs

![Graphs of sin, cos, and tan functions](image)

A.6 Trigonometry — special triangles

![Special triangles](image)

From the above pair of special triangles we have

\[
\begin{align*}
\sin \frac{\pi}{4} &= \frac{1}{\sqrt{2}} \\
\cos \frac{\pi}{4} &= \frac{1}{\sqrt{2}} \\
\tan \frac{\pi}{4} &= 1 \\
\sin \frac{\pi}{6} &= \frac{1}{2} \\
\cos \frac{\pi}{6} &= \frac{\sqrt{3}}{2} \\
\tan \frac{\pi}{6} &= \frac{1}{\sqrt{3}} \\
\sin \frac{\pi}{3} &= \frac{\sqrt{3}}{2} \\
\cos \frac{\pi}{3} &= \frac{1}{2} \\
\tan \frac{\pi}{3} &= \sqrt{3}
\end{align*}
\]

A.7 Trigonometry — simple identities

- Periodicity
  \[
  \sin(\theta + 2\pi) = \sin(\theta) \quad \cos(\theta + 2\pi) = \cos(\theta)
  \]

- Reflection
  \[
  \sin(-\theta) = -\sin(\theta) \quad \cos(-\theta) = \cos(\theta)
  \]

- Reflection around \(\pi/4\)
  \[
  \sin\left(\frac{\pi}{2} - \theta\right) = \cos\theta \quad \cos\left(\frac{\pi}{2} - \theta\right) = \sin\theta
  \]
- Reflection around $\pi/2$
  \[
  \sin(\pi - \theta) = \sin \theta \quad \cos(\pi - \theta) = -\cos \theta
  \]
- Rotation by $\pi$
  \[
  \sin(\theta + \pi) = -\sin \theta \quad \cos(\theta + \pi) = -\cos \theta
  \]
- Pythagoras
  \[
  \sin^2 \theta + \cos^2 \theta = 1
  \]

### A.8 Trigonometry — add and subtract angles

- Sine
  \[
  \sin(\alpha \pm \beta) = \sin(\alpha) \cos(\beta) \pm \cos(\alpha) \sin(\beta)
  \]
- Cosine
  \[
  \cos(\alpha \pm \beta) = \cos(\alpha) \cos(\beta) \mp \sin(\alpha) \sin(\beta)
  \]

### A.9 Inverse trigonometric functions

Some of you may not have studied inverse trigonometric functions in highschool, however we still expect you to know them by the end of the course.

<table>
<thead>
<tr>
<th></th>
<th>arcsin $x$</th>
<th>arccos $x$</th>
<th>arctan $x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain</td>
<td>$-1 \leq x \leq 1$</td>
<td>$-1 \leq x \leq 1$</td>
<td>all real numbers</td>
</tr>
<tr>
<td>Range</td>
<td>$-\frac{\pi}{2} \leq \arcsin x \leq \frac{\pi}{2}$</td>
<td>$0 \leq \arccos x \leq \pi$</td>
<td>$-\frac{\pi}{2} &lt; \arctan x &lt; \frac{\pi}{2}$</td>
</tr>
</tbody>
</table>

Since these functions are inverses of each other we have

\[
\begin{align*}
\arcsin(\sin \theta) &= \theta & \quad & -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2} \\
\arccos(\cos \theta) &= \theta & \quad & 0 \leq \theta \leq \pi \\
\arctan(\tan \theta) &= \theta & \quad & -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}
\end{align*}
\]
and also

\[
\begin{align*}
\sin(\arcsin x) &= x & -1 \leq x \leq 1 \\
\cos(\arccos x) &= x & -1 \leq x \leq 1 \\
\tan(\arctan x) &= x & \text{any real } x
\end{align*}
\]

### A.10 Areas

- **Area of a rectangle**
  \[A = bh\]

- **Area of a triangle**
  \[A = \frac{1}{2}bh = \frac{1}{2}ab \sin \theta\]

- **Area of a circle**
  \[A = \pi r^2\]

- **Area of an ellipse**
  \[A = \pi ab\]

### A.11 Volumes

- **Volume of a rectangular prism**
  \[V = lwh\]
- Volume of a cylinder

\[ V = \pi r^2 h \]

- Volume of a cone

\[ V = \frac{1}{3} \pi r^2 h \]

- Volume of a sphere

\[ V = \frac{4}{3} \pi r^3 \]

### A.12 Powers

In the following, \( x \) and \( y \) are arbitrary real numbers, and \( q \) is an arbitrary constant that is strictly bigger than zero.

- \( q^0 = 1 \)
- \( q^{x+y} = q^x q^y, \ q^{x-y} = \frac{q^x}{q^y} \)
- \( q^{-x} = \frac{1}{q^x} \)
- \( (q^x)^y = q^{xy} \)
- \( \lim_{x \to \infty} q^x = \infty, \ \lim_{x \to -\infty} q^x = 0 \) if \( q > 1 \)
- \( \lim_{x \to \infty} q^x = 0, \ \lim_{x \to -\infty} q^x = \infty \) if \( 0 < q < 1 \)
- The graph of \( 2^x \) is given below. The graph of \( q^x \), for any \( q > 1 \), is similar.
A.13 Logarithms

In the following, $x$ and $y$ are arbitrary real numbers that are strictly bigger than 0, and $p$ and $q$ are arbitrary constants that are strictly bigger than one.

- $q^{\log_q x} = x$, $\log_q (q^x) = x$
- $\log_q x = \frac{\log_p x}{\log_p q}$
- $\log_q 1 = 0$, $\log_q q = 1$
- $\log_q (xy) = \log_q x + \log_q y$
- $\log_q \left(\frac{x}{y}\right) = \log_q x - \log_q y$
- $\log_q \left(\frac{1}{y}\right) = -\log_q y$
- $\log_q (x^y) = y \log_q x$
- $\lim_{x \to \infty} \log_q x = \infty$, $\lim_{x \to 0} \log_q x = -\infty$
- The graph of $\log_{10} x$ is given below. The graph of $\log_q x$, for any $q > 1$, is similar.

A.14 Highschool material you should be able to derive

- Graphs of $\csc \theta$, $\sec \theta$ and $\cot \theta$:
- More Pythagoras

\[
\begin{align*}
\sin^2 \theta + \cos^2 \theta &= 1 & \text{divide by } \cos^2 \theta & \Rightarrow \tan^2 \theta + 1 = \sec^2 \theta \\
\sin^2 \theta + \cos^2 \theta &= 1 & \text{divide by } \sin^2 \theta & \Rightarrow 1 + \cot^2 \theta = \csc^2 \theta
\end{align*}
\]

- Sine — double angle (set \( \beta = \alpha \) in sine angle addition formula)

\[
\sin(2\alpha) = 2 \sin(\alpha) \cos(\alpha)
\]

- Cosine — double angle (set \( \beta = \alpha \) in cosine angle addition formula)

\[
\begin{align*}
\cos(2\alpha) &= \cos^2(\alpha) - \sin^2(\alpha) \\
&= 2 \cos^2(\alpha) - 1 && \text{(use } \sin^2(\alpha) = 1 - \cos^2(\alpha)\text{)} \\
&= 1 - 2 \sin^2(\alpha) && \text{(use } \cos^2(\alpha) = 1 - \sin^2(\alpha)\text{)}
\end{align*}
\]

- Composition of trigonometric and inverse trigonometric functions:

\[
\begin{align*}
\cos(\arcsin x) &= \sqrt{1 - x^2} \\
\sec(\arctan x) &= \sqrt{1 + x^2}
\end{align*}
\]

and similar expressions.

## A.15 Cartesian Coordinates

Each point in two dimensions may be labeled by two coordinates \((x, y)\) which specify the position of the point in some units with respect to some axes as in the figure below.

The set of all points in two dimensions is denoted \( \mathbb{R}^2 \). Observe that

- the distance from the point \((x, y)\) to the \(x\)-axis is \(|y|\)
- the distance from the point \((x, y)\) to the \(y\)-axis is \(|x|\)
- the distance from the point \((x, y)\) to the origin \((0, 0)\) is \(\sqrt{x^2 + y^2}\)

Similarly, each point in three dimensions may be labeled by three coordinates \((x, y, z)\), as in the two figures below.
The set of all points in three dimensions is denoted $\mathbb{R}^3$. The plane that contains, for example, the $x$– and $y$–axes is called the $xy$–plane.

- The $xy$–plane is the set of all points $(x, y, z)$ that obey $z = 0$.
- The $xz$–plane is the set of all points $(x, y, z)$ that obey $y = 0$.
- The $yz$–plane is the set of all points $(x, y, z)$ that obey $x = 0$.

More generally,

- The set of all points $(x, y, z)$ that obey $z = c$ is a plane that is parallel to the $xy$–plane and is a distance $|c|$ from it. If $c > 0$, the plane $z = c$ is above the $xy$–plane. If $c < 0$, the plane $z = c$ is below the $xy$–plane. We say that the plane $z = c$ is a signed distance $c$ from the $xy$–plane.
- The set of all points $(x, y, z)$ that obey $y = b$ is a plane that is parallel to the $xz$–plane and is a signed distance $b$ from it.
- The set of all points $(x, y, z)$ that obey $x = a$ is a plane that is parallel to the $yz$–plane and is a signed distance $a$ from it.

Observe that

- the distance from the point $(x, y, z)$ to the $xy$–plane is $|z|
- the distance from the point $(x, y, z)$ to the $xz$–plane is $|y|
- the distance from the point $(x, y, z)$ to the $yz$–plane is $|x|
- the distance from the point $(x, y, z)$ to the origin $(0, 0, 0)$ is $\sqrt{x^2 + y^2 + z^2}$
The distance from the point \((x, y, z)\) to the point \((x', y', z')\) is
\[
\sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}
\]
so that the equation of the sphere centered on \((1, 2, 3)\) with radius 4, that is, the set of all points \((x, y, z)\) whose distance from \((1, 2, 3)\) is 4, is
\[
(x - 1)^2 + (y - 2)^2 + (z - 3)^2 = 16
\]

### A.16 Roots of Polynomials

Being able to factor polynomials is a very important part of many of the computations in this course. Related to this is the process of finding roots (or zeros) of polynomials. That is, given a polynomial \(P(x)\), find all numbers \(r\) so that \(P(r) = 0\).

In the case of a quadratic \(P(x) = ax^2 + bx + c\), we can use the formula
\[
x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}
\]
The corresponding formulas for cubics and quartics\(^1\) are extremely cumbersome, and no such formula exists for polynomials of degree 5 and higher\(^2\).

Despite this there are many tricks\(^3\) for finding roots of polynomials that work well in some situations but not all. Here we describe approaches that will help you find integer and rational roots of polynomials that will work well on exams, quizzes and homework assignments.

Consider the quadratic equation \(x^2 - 5x + 6 = 0\). We could\(^4\) solve this using the quadratic formula
\[
x = \frac{5 \pm \sqrt{25 - 4 \times 1 \times 6}}{2} = \frac{5 \pm 1}{2} = 2, 3.
\]
Hence \(x^2 - 5x + 6\) has roots \(x = 2, 3\) and so it factors as \((x - 3)(x - 2)\). Notice\(^5\) that the numbers 2 and 3 divide the constant term of the polynomial, 6. This happens in general and forms the basis of our first trick.

**Trick 1 (A very useful trick).**

If \(r\) or \(-r\) is an integer root of a polynomial \(P(x) = a_nx^n + \cdots + a_1x + a_0\) with integer coefficients, then \(r\) is a factor of the constant term \(a_0\).

---

1. The method for cubics was developed in the 15th century by del Ferro, Cardano and Ferrari (Cardano’s student). Ferrari then went on to discover a formula for the roots of a quartic. His formula requires the solution of an associated cubic polynomial.
2. This is the famous Abel-Ruffini theorem.
3. There is actually a large body of mathematics devoted to developing methods for factoring polynomials. Polynomial factorisation is a fundamental problem for most computer algebra systems. The interested reader should make use of their favourite search engine to find out more.
4. We probably shouldn’t do it this way for such a simple polynomial, but for pedagogical purposes we do here.
5. Many of you may have been taught this approach in highschool.
Proof. If \( r \) is a root of the polynomial we know that \( P(r) = 0 \). Hence
\[
a_n \cdot r^n + \cdots + a_1 \cdot r + a_0 = 0
\]
If we isolate \( a_0 \) in this expression we get
\[
a_0 = -[a_n r^n + \cdots + a_1 r]
\]
We can see that \( r \) divides every term on the right-hand side. This means that the right-hand side is an integer times \( r \). Thus the left-hand side, being \( a_0 \), is an integer times \( r \), as required. The argument for when \(-r\) is a root is almost identical.

Let us put this observation to work.

**Example A.16.1**

Find the integer roots of \( P(x) = x^3 - x^2 + 2 \).

*Solution.*

- The constant term in this polynomial is 2.
- The only divisors of 2 are 1, 2. So the only candidates for integer roots are \( \pm 1, \pm 2 \).
- Trying each in turn
  \[
  P(1) = 2 \quad P(-1) = 0 \\
  P(2) = 6 \quad P(-2) = -10
  \]
- Thus the only integer root is \(-1\).

**Example A.16.2**

Find the integer roots of \( P(x) = 3x^3 + 8x^2 - 5x - 6 \).

*Solution.*

- The constant term is -6.
- The divisors of 6 are 1, 2, 3, 6. So the only candidates for integer roots are \( \pm 1, \pm 2, \pm 3, \pm 6 \).
- We try each in turn (it is tedious but not difficult):
  \[
  P(1) = 0 \quad P(-1) = 4 \\
  P(2) = 40 \quad P(-2) = 12 \\
  P(3) = 132 \quad P(-3) = 0 \\
  P(6) = 900 \quad P(-6) = -336
  \]
- Thus the only integer roots are 1 and -3.
We can generalise this approach in order to find rational roots. Consider the polynomial $6x^2 - x - 2$. We can find its zeros using the quadratic formula:

$$x = \frac{1 \pm \sqrt{1 + 48}}{12} = \frac{1 \pm 7}{12} = -\frac{1}{2}, \frac{2}{3}.$$ 

Notice now that the numerators, 1 and 2, both divide the constant term of the polynomial (being 2). Similarly, the denominators, 2 and 3, both divide the coefficient of the highest power of $x$ (being 6). This is quite general.

**Trick 2 (Another nice trick).**

If $b/d$ or $-b/d$ is a rational root in lowest terms (i.e. $b$ and $d$ are integers with no common factors) of a polynomial $Q(x) = a_nx^n + \cdots + a_1x + a_0$ with integer coefficients, then the numerator $b$ is a factor of the constant term $a_0$ and the denominator $d$ is a factor of $a_n$.

**Proof.** Since $b/d$ is a root of $P(x)$ we know that

$$a_n(b/d)^n + \cdots + a_1(b/d) + a_0 = 0$$

Multiply this equation through by $d^n$ to get

$$a_nb^n + \cdots + a_1bd^{n-1} + a_0d^n = 0$$

Move terms around to isolate $a_0d^n$:

$$a_0d^n = -[a_nb^n + \cdots + a_1bd^{n-1}]$$

Now every term on the right-hand side is some integer times $b$. Thus the left-hand side must also be an integer times $b$. We know that $d$ does not contain any factors of $b$, hence $a_0$ must be some integer times $b$ (as required).

Similarly we can isolate the term $a_nb^n$:

$$a_nb^n = -[a_{n-1}b^{n-1}d + \cdots + a_1bd^{n-1} + a_0d^n]$$

Now every term on the right-hand side is some integer times $d$. Thus the left-hand side must also be an integer times $d$. We know that $b$ does not contain any factors of $d$, hence $a_n$ must be some integer times $d$ (as required).

The argument when $-b/d$ is a root is nearly identical. 

We should put this to work:

**Example A.16.3**

$$P(x) = 2x^2 - x - 3.$$ 

**Solution.**
• The constant term in this polynomial is $3 = 1 \times 3$ and the coefficient of the highest power of $x$ is $2 = 1 \times 2$.

• Thus the only candidates for integer roots are $\pm 1, \pm 3$.

• By our newest trick, the only candidates for fractional roots are $\pm \frac{1}{2}, \pm \frac{3}{2}$.

• We try each in turn:

  $\begin{align*}
  P(1) &= -2 \\
  P(-1) &= 0 \\
  P(3) &= 12 \\
  P(-3) &= 18 \\
  P \left( \frac{1}{2} \right) &= -3 \\
  P \left( -\frac{1}{2} \right) &= -2 \\
  P \left( \frac{3}{2} \right) &= 0 \\
  P \left( -\frac{3}{2} \right) &= 3
  \end{align*}$

so the roots are $-1$ and $\frac{3}{2}$.

Example A.16.3

The tricks above help us to find integer and rational roots of polynomials. With a little extra work we can extend those methods to help us factor polynomials. Say we have a polynomial $P(x)$ of degree $p$ and have established that $r$ is one of its roots. That is, we know $P(r) = 0$. Then we can factor $(x - r)$ out from $P(x)$—it is always possible to find a polynomial $Q(x)$ of degree $p - 1$ so that

$P(x) = (x - r)Q(x)$

In sufficiently simple cases, you can probably do this factoring by inspection. For example, $P(x) = x^2 - 4$ has $r = 2$ as a root because $P(2) = 2^2 - 4 = 0$. In this case, $P(x) = (x - 2)(x + 2)$ so that $Q(x) = (x + 2)$. As another example, $P(x) = x^2 - 2x - 3$ has $r = -1$ as a root because $P(-1) = (-1)^2 - 2(-1) - 3 = 1 + 2 - 3 = 0$. In this case, $P(x) = (x + 1)(x - 3)$ so that $Q(x) = (x - 3)$.

For higher degree polynomials we need to use something more systematic—long division.

**Trick 3 (Long Division).**

Once you have found a root $r$ of a polynomial, even if you cannot factor $(x - r)$ out of the polynomial by inspection, you can find $Q(x)$ by dividing $P(x)$ by $x - r$, using the long division algorithm you learned in school, but with 10 replaced by $x$.

---

6 Again, this is a little tedious, but not difficult. It’s actually pretty easy to code up for a computer to do. Modern polynomial factoring algorithms do more sophisticated things, but these are a pretty good way to start.
Example A.16.4

Factor \( P(x) = x^3 - x^2 + 2. \)

Solution.

- We can go hunting for integer roots of the polynomial by looking at the divisors of the constant term. This tells us to try \( x = \pm 1, \pm 2. \)

- A quick computation shows that \( P(-1) = 0 \) while \( P(1), P(-2), P(2) \neq 0. \) Hence \( x = -1 \) is a root of the polynomial and so \( x + 1 \) must be a factor.

- So we divide \( \frac{x^3 - x^2 + 2}{x + 1}. \) The first term, \( x^2, \) in the quotient is chosen so that when you multiply it by the denominator, \( x^2(x + 1) = x^3 + x^2, \) the leading term, \( x^3, \) matches the leading term in the numerator, \( x^3 - x^2 + 2, \) exactly.

\[
x + 1 \left[ \frac{x^2}{x^3 + x^2} + \frac{2}{x + 1} \right] \rightarrow x^2(x + 1)
\]

- When you subtract \( x^2(x + 1) = x^3 + x^2 \) from the numerator \( x^3 - x^2 + 2 \) you get the remainder \( -2x^2 + 2. \) Just like in public school, the \( 2 \) is not normally “brought down” until it is actually needed.

\[
x + 1 \left[ \frac{x^2}{x^3 + x^2} + \frac{2}{x + 1} \right] \rightarrow x^2(x + 1)
\]

- The next term, \( -2x, \) in the quotient is chosen so that when you multiply it by the denominator, \( -2x(x + 1) = -2x^2 - 2x, \) the leading term \( -2x^2 \) matches the leading term in the remainder exactly.

\[
x + 1 \left[ \frac{x^2 - 2x}{x^3 + x^2} - \frac{2}{x + 1} \right] \rightarrow -2x(x + 1)
\]

And so on.

\[
x + 1 \left[ \frac{x^2 - 2x + 2}{x^3 + x^2} \right] \rightarrow 2(x + 1)
\]

7 This is a standard part of most highschool mathematics curricula, but perhaps not all. You should revise this carefully.
• Note that we finally end up with a remainder 0. A nonzero remainder would have signalled a computational error, since we know that the denominator \( x - (-1) \) must divide the numerator \( x^3 - x^2 + 2 \) exactly.

• We conclude that

\[
(x + 1)(x^2 - 2x + 2) = x^3 - x^2 + 2
\]

To check this, just multiply out the left hand side explicitly.

• Applying the high school quadratic root formula \( \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \) to \( x^2 - 2x + 2 \) tells us that it has no real roots and that we cannot factor it further.  

We finish by describing an alternative to long division. The approach is roughly equivalent, but is perhaps more straightforward at the expense of requiring more algebra.

\[\begin{align*}
\text{Example A.16.4} \\
\text{Factor } P(x) = x^3 - x^2 + 2, \text{ again.}
\end{align*}\]

\[\text{Solution.} \text{ Let us do this again but avoid long division.}
\]

• From the previous example, we know that \( \frac{x^3 - x^2 + 2}{x + 1} \) must be a polynomial (since \(-1\) is a root of the numerator) of degree 2. So write

\[
\frac{x^3 - x^2 + 2}{x + 1} = ax^2 + bx + c
\]

for some, as yet unknown, coefficients \( a, b \) and \( c \).

• Cross multiplying and simplifying gives us

\[
x^3 - x^2 + 2 = (ax^2 + bx + c)(x + 1) = ax^3 + (a + b)x^2 + (b + c)x + c
\]

• Now matching coefficients of the various powers of \( x \) on the left and right hand sides

- coefficient of \( x^3 \): \( a = 1 \)
- coefficient of \( x^2 \): \( a + b = -1 \)
- coefficient of \( x^1 \): \( b + c = 0 \)
- coefficient of \( x^0 \): \( c = 2 \)

• This gives us a system of equations that we can solve quite directly. Indeed it tells us immediately that that \( a = 1 \) and \( c = 2 \). Subbing \( a = 1 \) into \( a + b = -1 \) tells us that \( 1 + b = -1 \) and hence \( b = -2 \).
Thus

\[ x^3 - x^2 + 2 = (x + 1)(x^2 - 2x + 2). \]