SFU DEPARTMENT OF MATHEMATICS

CALCULUS
Early Transcendentals

Integral & Multi-Variable Calculus for Social Sciences

SFU DEPARTMENT OF MATHEMATICS

 Examples, in particular applications, have been added throughout the textbook to further illustrate concepts and to reflect social science calculus content; exercises have been added throughout the textbook to provide more extensive practice, and information throughout the book, as applicable, has been revised to reflect social science calculus content and modified to visually align definitions, theorems and guidelines established in the textbook.
The following additions have been made to these chapters:

Chapter 1:

- Antiderivatives

Chapter 2:

- Partial Fraction Method

Chapter 3:

- Business and Economics Applications

Chapter 4:

- Probability: One Random Variable, Two Random Variables

Chapter 5:

- Classifying Differential Equations
- Simple Growth and Decay Model, Logistic Growth Model
- Slope Fields

The following deletions have been made: Review, Functions, Limits, Derivatives, Applications of Derivatives, Selected Applications of Integration (Distance, Velocity, Acceleration, Work, Centre of Mass, Arc Length, Surface Area), Polar Coordinates, Parametric Equations, Three Dimensions, Partial Differentiation, Vector Functions, and Vector Calculus.

SFU DEPARTMENT OF MATHEMATICS

Acknowledgements

The production of this textbook and supplementary material was made possible under an SFU Open Educational Resources Grant and matching funds from the Department of Mathematics. The authors acknowledge the helpful advice received from the following staff at SFU:

Hope Power, Teaching & Learning Librarian, SFU

Jennifer Zerkee, Copyright Specialist, SFU Copyright Office

Cindy Xin, Educational Consultant, SFU Teaching and Learning Centre.

We are indebted to John Stockie for his careful proofreading of the text, but moreover, for providing constructive criticism that further helped shape the adaption into a cohesive student-oriented calculus text.

OER Support: Corrections and Suggestions

Please support OER. In an effort to improve the content of this textbook, contact Petra Menz at pmenz@sfu.ca with your suggestions for improvements, new content, or errata.

Dedication

To my son Eli, so that his access to learning remains open as he unfolds his wings to explore life. May his roots be strong and anchor him in all of his pursuits.

Petra Menz

Contents

Introduction

This course is designed for students specializing in business or the social sciences. Topics include integration; a variety of techniques of integration; applications of integration such as average value, area, volume using Disk and Shell Methods, surplus in consumption and production, continuous money flow, double and triple integration, and probability; differential equations; and sequences and series leading up to Taylor's Theorem.

The following *Recommendations for Success in Mathematics* are excerpts taken from the same named document published by Petra Menz in order to provide strategies grouped into categories to all students who are thinking about their well-being, learning, and goals, and who want to be successful academically.

How to Take Lecture Notes:

Listen to the Instructor, who

- explains the concepts;
- draws connections;
- demonstrates examples;
- emphasizes material.

Copy the presented lecture material

- by arriving to the lecture prepared;
- using telegraphic writing, i.e. packing as much information into the smallest possible number of words/ symbols (do you really need to copy all the algebraic/manipulative steps?).

Mark up your notes immediately while listening and copying using a system such as offered here:

! pay attention (possible exam material)

? confusing (read course notes or visit ACW)

-> practice (using course notes and online assignments)

__ underline/highlight key concepts

Habits of a Successful Student: for detailed description of each item see the document *Recommendations for Success in Mathematics*

- Acts responsibly
- Sets goals
- Is reflective
- Is inquisitive
- Can communicate
- Enjoys learning
- Is resourceful
- Is organized
- Manages time effectively
- Is involved

Problem Solving Strategies

The emphasis in this course is on problems–doing calculations and applications. To master problem solving one needs a tremendous amount of practice doing problems. The more problems you do the better you will be at doing them, as patterns will start to emerge in both the problems and in successful approaches to them. You will learn quickly and effectively if you devote some time to doing problems every day.

Typically the most difficult problems are applications, since they require some effort before you can begin calculating. Here are some pointers for doing applications:

1. Carefully read each problem twice before writing anything.
2. Assign letters to quantities that are described only in words; draw a diagram if appropriate.
3. Decide which letters are constants (invariants) and which are variables (variants). A letter stands for a constant if its value remains the same throughout the problem. A letter stands for a variable if its value varies throughout the problem.
4. Using mathematical notation, write down what you know and then write down what you want to find.
5. Decide what category of problem it is (this might be obvious if the problem comes at the end of a particular chapter, but will not necessarily be so obvious if it comes on an exam covering several chapters).
6. Double check each step as you go along; don't wait until the end to check your work.
7. Use common sense; if an answer is out of the range of practical possibilities, then check your work to see where you went wrong

1. Integration

1.1 Antiderivatives

You have probably taken a course in *Differential Calculus*, where you have solved problems of the following nature:

- Given $f(x) = \sin(x) + 3x^5 - 12$, find $f'(x)$.
- Given the position function of an object, determine its velocity function.
- Given the demand function, calculate the elasticity of demand.

To solve any of these problems we need the concept of the **derivative**, which provides us with information about the rate of change of the quantity involved that leads to the solution. In other words, *Differential Calculus* allows us to solve problems that are concerned with finding the rate of change of one quantity with respect to another quantity.

The next three chapters are based on the idea of the **antiderivative**, which basically helps us solve problems that are the *reverse* of the above problems such as

- Given $f'(x) = \sin(x) + 3x^5 - 12$, find $f(x)$.
- Given the velocity function of an object, determine its position function.
- How much do consumers benefit by purchasing some manufactured goods at the price determined by supply and demand?

We will develop tools for finding the antiderivative, which is the process of **antidifferentiation** also known as **integration**. Hence, these kinds of problems fall under the topic of *Integral Calculus*. In summary, *Integral Calculus* allows us to solve problems, where the rate of change of one quantity is given with respect to another quantity and we are concerned with finding the relationship between these two quantities.

Definition 1.1: Antiderivative

A function F is an **antiderivative** *of f on an interval I if*

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

for all x in I.

1.2 Displacement and Area

We will now delve into two examples, one concerning displacement and the other area, to explore how the antidifferentiation process should unfold.

Example 1.2: Object Moving in a Straight Line

An object moves in a straight line so that its speed at time t is given by $v(t) = 3t$ in, say, cm/sec. If the object is at position 10 on the straight line when $t = 0$, where is the object at any time t?

Solution. There are two reasonable ways to approach this problem.

Method 1: If $s(t)$ is the position of the object at time t, we know that $s'(t) = v(t)$. Based on our knowledge of derivatives, we therefore know that

$$s(t) = 3t^2/2 + k,$$

and because $s(0) = 10$ we easily discover that $k = 10$, so

$$s(t) = 3t^2/2 + 10.$$

For example, at $t = 1$ the object is at position $3/2 + 10 = 11.5$ cm.

This is certainly the easiest way to deal with this problem. Not all similar problems are so easy, as we will see; the second approach to the problem is more difficult but also more general.

Method 2: We start by considering how we might approximate a solution. We know that at $t = 0$ the object is at position 10. How might we approximate its position at, say, $t = 1$? We know that the speed of the object at time $t = 0$ is 0; if its speed were constant then in the first second the object would not move and its position would still be 10 when $t = 1$. In fact, the object will not be too far from 10 at $t = 1$, but certainly we can do better.

Let's look at the times 0.1, 0.2, 0.3, ..., 1.0, and try approximating the location of the object at each, by supposing that during each tenth of a second the object is going at a constant speed. Since the object initially has speed 0, we again suppose it maintains this speed, but only for a tenth of second; during that time the object would not move. During the tenth of a second from $t = 0.1$ to $t = 0.2$, we suppose that the object is travelling at 0.3 cm/sec, namely, its actual speed at $t = 0.1$. In this case the object would travel $(0.3)(0.1) = 0.03$ centimetres: 0.3 cm/sec times 0.1 seconds. Similarly, between $t = 0.2$ and $t = 0.3$ the object would travel $(0.6)(0.1) = 0.06$ centimetres. Continuing, we get as an approximation that the object travels

$$(0.0)(0.1) + (0.3)(0.1) + (0.6)(0.1) + \cdots + (2.7)(0.1) = 1.35$$

centimetres, ending up at position 11.35 cm. This is a better approximation than 10, certainly, but is still just an approximation. (We know in fact that the object ends up at position 11.5, because we've already done the problem using the first approach.)

Presumably, we will get a better approximation if we divide the time into one hundred intervals of a hundredth of a second each, and repeat the process:

$$(0.0)(0.01) + (0.03)(0.01) + (0.06)(0.01) + \cdots + (2.97)(0.01) = 1.485.$$

We thus approximate the position as 11.485. Since we know the exact answer, we can see that this is much closer, but if we did not already know the answer, we wouldn't really know how close.

We can keep this up, but we'll never really know the exact answer if we simply compute more and more examples. Let's instead look at a "typical" approximation. Suppose we divide the time into n equal intervals, and imagine that on each of these the object travels at a constant speed. Over the first time interval we approximate the distance travelled as $(0.0)(1/n) = 0$, as before. During the second time interval, from $t = 1/n$ to $t = 2/n$, the object travels approximately $3(1/n)(1/n) = 3/n^2$ centimetres. During time interval number i, the object travels approximately $(3(i-1)/n)(1/n) = 3(i-1)/n^2$ centimetres, that is, its speed at time $(i-1)/n$, $3(i-1)/n$, times the length of time interval number i, $1/n$. Adding these up as before, we approximate the distance travelled as

$$(0)\frac{1}{n} + 3\frac{1}{n^2} + 3(2)\frac{1}{n^2} + 3(3)\frac{1}{n^2} + \cdots + 3(n-1)\frac{1}{n^2}$$

centimetres. What can we say about this? At first it looks rather less useful than the concrete calculations we've already done, but in fact a bit of algebra reveals it to be much more useful. We can factor out a 3 and $1/n^2$ to get

$$\frac{3}{n^2}(0+1+2+3+\cdots+(n-1)),$$

that is, $3/n^2$ times the sum of the first $n-1$ positive integers.

Now we make use of a fact you may have run across before, Gauss's Equation:

$$1+2+3+\cdots+k = \frac{k(k+1)}{2}.$$

In our case we're interested in $k = n-1$, so

$$1+2+3+\cdots+(n-1) = \frac{(n-1)(n)}{2} = \frac{n^2-n}{2}.$$

This simplifies the approximate distance travelled to

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

Now this is quite easy to understand: as n gets larger and larger this approximation gets closer and closer to $(3/2)(1-0) = 3/2$, so that $3/2$ is the exact distance traveled during one second, and the final position is 11.5.

So for $t = 1$, at least, this rather cumbersome approach gives the same answer as the first approach. But really there's nothing special about $t = 1$; let's just call it t instead. In this case the approximate distance traveled during time interval number i is $3(i-1)(t/n)(t/n) = 3(i-1)t^2/n^2$, that is, speed $3(i-1)(t/n)$ times time t/n, and the total distance traveled is approximately

$$(0)\frac{t}{n} + 3(1)\frac{t^2}{n^2} + 3(2)\frac{t^2}{n^2} + 3(3)\frac{t^2}{n^2} + \cdots + 3(n-1)\frac{t^2}{n^2}.$$

As before we can simplify this to

$$\frac{3t^2}{n^2}(0+1+2+\cdots+(n-1)) = \frac{3t^2}{n^2}\frac{n^2-n}{2} = \frac{3}{2}t^2\left(1-\frac{1}{n}\right).$$

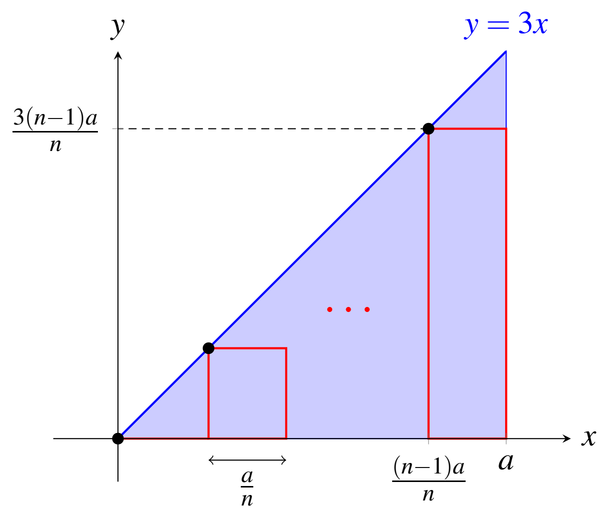
In the limit, as n gets larger, this gets closer and closer to $(3/2)t^2$ and the approximated position of the object gets closer and closer to $(3/2)t^2+10$, so the actual position is $(3/2)t^2+10$, exactly the answer given by the first method to the problem.

♣

Example 1.3: Area under the Line

Find the area under the curve $y=3x$ between $x=0$ and any positive value $x=a$.

Solution. There is here no obvious analogue to the first method in the previous example, but the second method works fine. (Since the function $y=3x$ is so simple, there is another method that works here, but it is even more limited in potential application than is method number one.) How might we approximate the desired area? We know how to compute areas of rectangles, so we approximate the area by rectangles as shown below.



Jumping straight to the general case, suppose we divide the interval between 0 and a into n equal subintervals, and use a rectangle above each subinterval to approximate the area under the curve. There are many ways we might do this, but let's use the height of the curve at the left endpoint of the subinterval as the height of the rectangle. The height of rectangle number i is then $3(i-1)(a/n)$, the width is a/n, and the area is $3(i-1)(a^2/n^2)$. The total area of the rectangles is

$$(0)\frac{a}{n}+3(1)\frac{a^2}{n^2}+3(2)\frac{a^2}{n^2}+3(3)\frac{a^2}{n^2}+\cdots+3(n-1)\frac{a^2}{n^2}.$$

By factoring out $3a^2/n^2$ this simplifies to

$$\frac{3a^2}{n^2}(0+1+2+\cdots+(n-1))=\frac{3a^2}{n^2}\frac{n^2-n}{2}=\frac{3}{2}a^2\left(1-\frac{1}{n}\right).$$

As n gets larger this gets closer and closer to $3a^2/2$, which must therefore be the true area under the curve.

♣

What you will have noticed, of course, is that while the problem in the second example appears to be much different than the problem in the first example, and while the easy approach to problem one does not appear to apply to problem two, the "approximation" approach works in both, and moreover the *calculations are identical.* As we will see, there are many, many problems that appear much different on the surface but turn out to be the same as these problems, in the sense that when we try to approximate solutions we end up with mathematics that looks like the two examples, though of course the function involved will not always be so simple.

Even better, we now see that while the second problem did not appear to be amenable to approach one, it can in fact be solved in the same way. The reasoning is this: we know that problem one can be solved easily by finding a function whose derivative is $3t$. We also know that mathematically the two problems are the same, because both can be solved by taking a limit of a sum, and the sums are identical. Therefore, we don't really need to compute the limit of either sum because we know that we will get the same answer by computing a function with the derivative $3t$ or, which is the same thing, $3x$.

It's true that the first problem had the added complication of the "10", and we certainly need to be able to deal with such minor variations, but that turns out to be quite simple. The lesson then is this: whenever we can solve a problem by taking the limit of a sum of a certain form, instead of computing the (often nasty) limit we can find a new function with a certain derivative.

Exercises for Section 1.2

Exercise 1.2.1 *Suppose an object moves in a straight line so that its speed at time t is given by* $v(t) = 2t+2$*, and that at* $t = 1$ *the object is at position 5. Find the position of the object at* $t = 2$*.*

Exercise 1.2.2 *Suppose an object moves in a straight line so that its speed at time t is given by* $v(t) = t^2+2$*, and that at* $t = 0$ *the object is at position 5. Find the position of the object at* $t = 2$*.*

1.3 Riemann Sums

A fundamental calculus technique is to first answer a given problem with an approximation, then refine that approximation to make it better, then use limits in the refining process to find the exact answer. That is exactly what we will do here to develop a technique to find the area of more complicated regions.

Consider the region given in Figure 1.1, which is the area under $y = 4x - x^2$ on $[0,4]$. What is the signed area of this region when the area above the x-axis is positive and below negative?

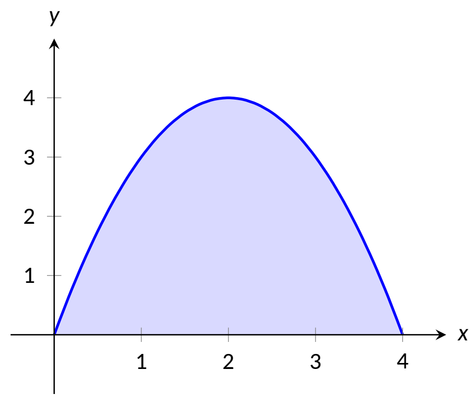


Figure 1.1: $f(x) = 4x - x^2$

We start by approximating. We can surround the region with a rectangle with height and width of 4 and find the area is approximately 16 square units. This is obviously an *over–approximation*; we are including area in the rectangle that is not under the parabola. How can we refine our approximation to make it better? The key to this section is this answer: *use more rectangles.*

Let's use four rectangles of equal width of 1. This partitions the interval $[0,4]$ into four *subintervals*, $[0,1]$, $[1,2]$, $[2,3]$ and $[3,4]$. On each subinterval we will draw a rectangle.

There are three common ways to determine the height of these rectangles: the **Left Hand Rule**, the **Right Hand Rule**, and the **Midpoint Rule**:

- The **Left Hand Rule** says to evaluate the function at the left-hand endpoint of the subinterval and make the rectangle that height. In Figure 1.2, the rectangle labelled "LHR" is drawn on the interval $[2,3]$ with a height determined by the Left Hand Rule, namely $f(2) = 4$.

- The **Right Hand Rule** says the opposite: on each subinterval, evaluate the function at the right endpoint and make the rectangle that height. In Figure 1.2, the rectangle labelled "RHR" is drawn on the interval $[0,1]$ with a height determined by the Right Hand Rule, namely $f(1) = 3$.

- The **Midpoint Rule** says that on each subinterval, evaluate the function at the midpoint and make the rectangle that height. In Figure 1.2, the rectangle labelled "MPR" is drawn on the interval $[1,2]$ with a height determined by the Midpoint Rule, namely $f(1.5) = 3.75$.

- These are the three most common rules for determining the heights of approximating rectangles, but we are not forced to use one of these three methods. In Figure 1.2, the rectangle labelled "other" is drawn on the interval $[3,4]$ with a height determined by choosing a random x-value on the interval $[3,4]$. The chosen x-value is 3.54, which yields a height of $f(3.54)$.

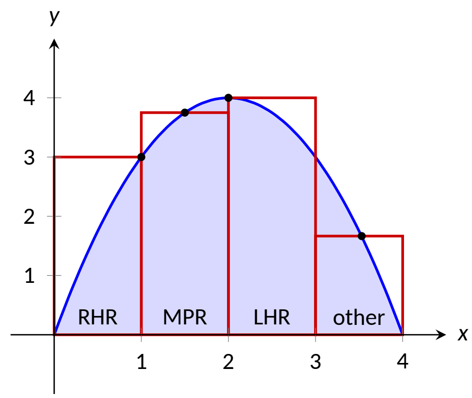


Figure 1.2: Approximating area using rectangles

The following example will approximate the area under $f(x) = 4x - x^2$ using these rules.

Example 1.4: Using the Left Hand, Right Hand and Midpoint Rules

Approximate the area under $f(x) = 4x - x^2$ on the interval $[0,4]$ using the Left Hand Rule, the Right Hand Rule, and the Midpoint Rule, using four equally spaced subintervals.

Solution. We break the interval $[0,4]$ into four subintervals as before. In Figure 1.3 we see four rectangles drawn on $f(x) = 4x - x^2$ using the Left Hand Rule. (The areas of the rectangles are given in each figure.)

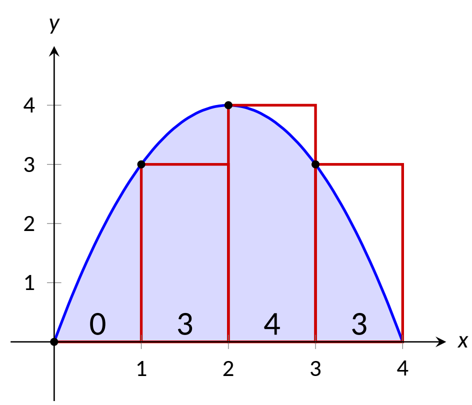


Figure 1.3: Approximating area using the Left Hand Rule

Note how in the first subinterval, $[0,1]$, the rectangle has height $f(0) = 0$. We add up the areas of each rectangle (height$\times$ width) for our Left Hand Rule approximation:

$$\begin{aligned} f(0)\cdot 1 + f(1)\cdot 1 + f(2)\cdot 1 + f(3)\cdot 1 = \\ 0+3+4+3 = 10. \end{aligned}$$

Figure 1.4 shows four rectangles drawn under $f(x)$ using the Right Hand Rule; note how the $[3,4]$ subinterval has a rectangle of height 0.

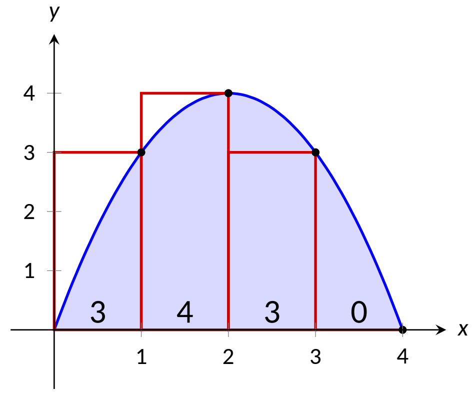


Figure 1.4: Approximating area using the Right Hand Rule,

In this figure, these rectangles seem to be the mirror image of those found in Figure 1.3. (This is because of the symmetry of our shaded region.) Our approximation gives the same answer as before, though calculated a different way:

$$f(1)\cdot 1+f(2)\cdot 1+f(3)\cdot 1+f(4)\cdot 1 = \\ 3+4+3+0 = 10.$$

Figure 1.5 shows four rectangles drawn under $f(x)$ using the Midpoint Rule.

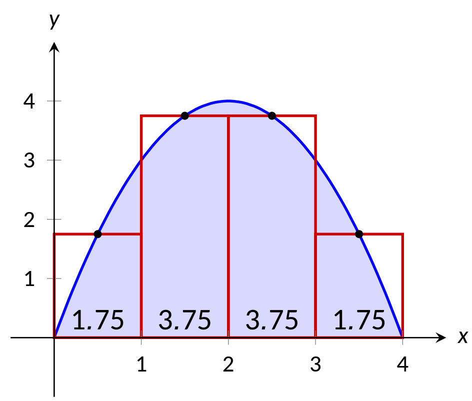


Figure 1.5: Approximating area using the Midpoint Rule

This gives an approximation of the area as:

$$f(0.5)\cdot 1+f(1.5)\cdot 1+f(2.5)\cdot 1+f(3.5)\cdot 1 = \\ 1.75+3.75+3.75+1.75 = 11.$$

Our three methods provide two approximations of the area under $f(x) = 4x - x^2$: 10 and 11. ♣

It is hard to tell at this moment which is a better approximation: 10 or 11? We can continue to refine our approximation by using more rectangles. The notation can become unwieldy, though, as we add up longer and longer lists of numbers. We introduce **summation notation** (also called **sigma notation**) to solve this problem.

Definition 1.5: Sigma Notation

Given the sum $a_1 + a_2 + a_3 + \cdots + a_{n-1} + a_n$, we use **sigma notation** *to write the sum in the compact form*

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + a_3 + \cdots + a_{n-1} + a_n,$$

where

$\sum_{i=1}^{n} a_i$ *is read "the sum as i goes from 1 to n of a_i",*

$\sum$ *is the Greek letter sigma and used as the* **summation symbol**,

the variable i is called the **index** *and takes on only integer values,*

the index i starts at $i = 1$ and ends at $i = n$, and

a_i represents the formula for the i-th term.

Note:

1. The index is often denoted by i, k or n and must be written below the summation symbol.
2. Do not mix the index up with the end-value of the index that must be written above the summation symbol.
3. The index can start at any integer, but often we write the sum so that the index starts at 0 or 1.

Let's practice using this notation.

Example 1.6: Using Summation Notation

Let the numbers $\{a_i\}$ be defined as $a_i = 2i - 1$ for integers i, where $i \geq 1$. So $a_1 = 1$, $a_2 = 3$, $a_3 = 5$, etc. (The output is the positive odd integers). Evaluate the following summations:

(a) $\sum_{i=1}^{6} a_i$ *(b)* $\sum_{i=3}^{7} (3a_i - 4)$ *(c)* $\sum_{i=1}^{4} (a_i)^2$

Solution.

(a)
$$\begin{aligned}\sum_{i=1}^{6} a_i &= a_1 + a_2 + a_3 + a_4 + a_5 + a_6 \\ &= 1 + 3 + 5 + 7 + 9 + 11 \\ &= 36.\end{aligned}$$

(b) Note the starting value is different than 1:

$$\sum_{i=3}^{7} (3a_i - 4) = (3a_3 - 4) + (3a_4 - 4) + (3a_5 - 4) + (3a_6 - 4) + (3a_7 - 4)$$

$$= 11 + 17 + 23 + 29 + 35$$
$$= 115.$$

(c) $\displaystyle\sum_{i=1}^{4}(a_i)^2 = (a_1)^2 + (a_2)^2 + (a_3)^2 + (a_4)^2$

$$= 1^2 + 3^2 + 5^2 + 7^2$$
$$= 84$$

♣

The following theorems give some properties and formulas of summations that allow us to work with them without writing individual terms. Examples will follow.

Theorem 1.7: Summation Properties

For c constant:

1. $\displaystyle\sum_{i=1}^{n} c = c \cdot n$
2. $\displaystyle\sum_{i=m}^{n}(a_i \pm b_i) = \sum_{i=m}^{n} a_i \pm \sum_{i=m}^{n} b_i$
3. $\displaystyle\sum_{i=m}^{n} c \cdot a_i = c \cdot \sum_{i=m}^{n} a_i$
4. $\displaystyle\sum_{i=m}^{j} a_i + \sum_{i=j+1}^{n} a_i = \sum_{i=m}^{n} a_i$

Theorem 1.8: Summation Formulas

1. $\displaystyle\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$
2. $\displaystyle\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$
3. $\displaystyle\sum_{i=1}^{n} i^3 = \left(\frac{n(n+1)}{2}\right)^2$

Example 1.9: Evaluating Summations

Evaluate $\displaystyle\sum_{i=1}^{6}(2i-1)$.

Solution.

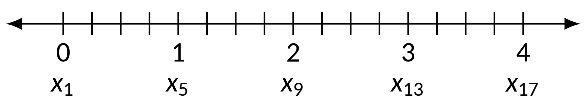
$$\begin{aligned}\sum_{i=1}^{6}(2i-1) &= \sum_{i=1}^{6}2i-\sum_{i=1}^{6}(1) \quad \text{(using Summation Property 2)}\\ &= \left(2\sum_{i=1}^{6}i\right)-6 \quad \text{(using Summation Properties 1 and 3)}\\ &= 2\frac{6(6+1)}{2}-6 \quad \text{(using Summation Formula 1)}\\ &= 42-6=36\end{aligned}$$

We obtained the same answer without writing out all six terms. When dealing with small values of n, it may be faster to write the terms out by hand. However, Theorems 1.7 and 1.8 are incredibly important when dealing with large sums as we'll soon see. ♣

Example 1.10: Creating Right Hand, Left Hand and Midpoint Rule Formulas

Suppose a continuous function $y = f(x)$ is defined on the interval $[0,4]$. Create the summation formulas for approximating the area of f on the given interval using the Right Hand, Left Hand and Midpoint Rules.

Solution. We will do some careful preparation. We start with a number line where $[0,4]$ is divided into sixteen equally spaced subintervals with partition $P = \{x_1, x_2, \ldots, x_{17}\}$.



We denote 0 as x_1; we have marked the values of x_5, x_9, x_{13} and x_{17}. We could mark them all, but the figure would get crowded. While it is easy to figure that $x_{10} = 2.25$, in general, we want a method of determining the value of x_i without consulting the figure. Consider:

$$\begin{array}{rcl} x_i & = & x_1+(i-1)\Delta x \\ \textit{where} & & \\ x_1 & : & \text{starting value} \\ (i-1) & : & \text{number of subintervals between } x_1 \text{ and } x_i \\ \Delta x & : & \text{subinterval width} \end{array}$$

So $x_{10} = x_1 + 9(4/16) = 2.25$.

If we had partitioned $[0,4]$ into 100 equally spaced subintervals with partition $P = \{x_1, x_2, \ldots, x_{101}\}$, each subinterval would have length $\Delta x = 4/100 = 0.04$. We could compute x_{32} as

$$x_{32} = 0+31(4/100) = 1.24.$$

(That was far faster than creating a sketch first.)

Given *any* subdivision of $[0,4]$, the first subinterval is $[x_1, x_2]$; the second is $[x_2, x_3]$; the i^{th} subinterval is $[x_i, x_{i+1}]$. Now recall our work in Example 1.4 and the Figures 1.3, 1.4 and 1.5.

- When using the Left Hand Rule, the height of the i^{th} rectangle will be $f(x_i)$.
- When using the Right Hand Rule, the height of the i^{th} rectangle will be $f(x_{i+1})$.
- When using the Midpoint Rule, the height of the i^{th} rectangle will be $f\left(\frac{x_i+x_{i+1}}{2}\right)$.

Thus approximating the area under f on $[0,4]$ with sixteen equally spaced subintervals can be expressed as follows, where $\Delta x = 4/16 = 1/4$:

Left Hand Rule: $\displaystyle\sum_{i=1}^{16} f(x_i)\cdot\frac{1}{4}$

Right Hand Rule: $\displaystyle\sum_{i=1}^{16} f(x_{i+1})\cdot\frac{1}{4}$

Midpoint Rule: $\displaystyle\sum_{i=1}^{16} f\left(\frac{x_i+x_{i+1}}{2}\right)\cdot\frac{1}{4}$

♣

We use these formulas in the following example.

Example 1.11: Approximating Area Using Sums

Approximate the area under $f(x) = 4x - x^2$ on $[0,4]$ using the Right Hand Rule and summation formulas with sixteen and 1000 equally spaced intervals.

Solution. Using sixteen equally spaced intervals and the Right Hand Rule, we can approximate the area as

$$\sum_{i=1}^{16} f(x_{i+1})\Delta x.$$

We have $\Delta x = 4/16 = 0.25$. Since $x_i = 0 + (i-1)\Delta x$, we have

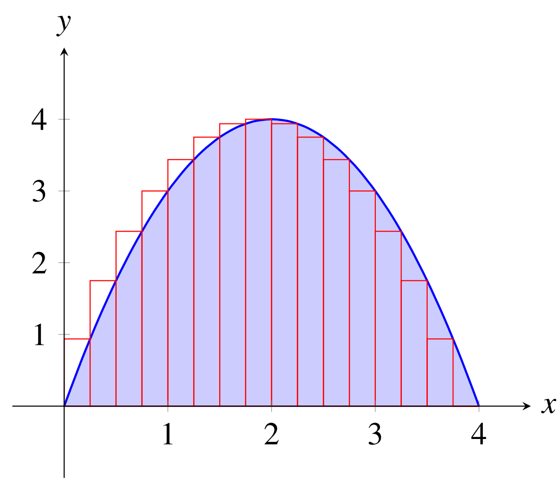
$$x_{i+1} = 0 + \big((i+1)-1\big)\Delta x = i\Delta x$$

Using the summation formulas, consider:

$$\begin{aligned}
\sum_{i=1}^{16} f(x_{i+1})\Delta x &= \sum_{i=1}^{16} f(i\Delta x)\Delta x \\
&= \sum_{i=1}^{16} \big(4i\Delta x - (i\Delta x)^2\big)\Delta x \\
&= \sum_{i=1}^{16} (4i\Delta x^2 - i^2\Delta x^3) \\
&= (4\Delta x^2)\sum_{i=1}^{16} i - \Delta x^3 \sum_{i=1}^{16} i^2 \qquad (1.1)
\end{aligned}$$

$$\begin{aligned}
&= (4\Delta x^2)\frac{16 \cdot 17}{2} - \Delta x^3 \frac{16(17)(33)}{6} \\
&= 4 \cdot 0.25^2 \cdot 136 - 0.25^3 \cdot 1496 \\
&= 10.625
\end{aligned}$$

We were able to sum up the areas of sixteen rectangles with very little computation. The function and the sixteen rectangles are graphed below. While some rectangles over–approximate the area, other under–approximate the area (by about the same amount). Thus our approximate area of 10.625 is likely a fairly good approximation.



Notice Equation (1.1); by changing the 16's to 1,000's (and appropriately changing the value of Δx), we can use that equation to sum up 1000 rectangles!

We do so here, skipping from the original summand to the equivalent of Equation (1.1) to save space. Note that $\Delta x = 4/1000 = 0.004$.

$$\begin{aligned}
\sum_{i=1}^{1000} f(x_{i+1})\Delta x &= (4\Delta x^2)\sum_{i=1}^{1000} i - \Delta x^3 \sum_{i=1}^{1000} i^2 \\
&= (4\Delta x^2)\frac{1000 \cdot 1001}{2} - \Delta x^3 \frac{1000(1001)(2001)}{6} \\
&= 4 \cdot 0.004^2 \cdot 500500 - 0.004^3 \cdot 333,833,500 \\
&= 10.666656
\end{aligned}$$

Using many, many rectangles, we have a likely good approximation of the area under $f(x) = 4x - x^2$ of ≈ 10.666656. ♣

Before the above example, we stated the summations for the Left Hand, Right Hand and Midpoint Rules in Example 1.10. Each had the same basic structure, which was:

1. Each rectangle has the same width, which we referred to as Δx, and
2. Each rectangle's height is determined by evaluating $f(x)$ at a particular point in each subinterval. For instance, the Left Hand Rule states that each rectangle's height is determined by evaluating $f(x)$ at the left hand endpoint of the subinterval the rectangle lives on.

One could partition an interval $[a,b]$ with subintervals that did not have the same width. We refer to the length of the first subinterval as Δx_1, the length of the second subinterval as Δx_2, and so on, giving the length of the i^{th} subinterval as Δx_i. Also, one could determine each rectangle's height by evaluating $f(x)$ at *any* point in the i^{th} subinterval. We refer to the point picked in the first subinterval as c_1, the point picked in the second subinterval as c_2, and so on, with c_i representing the point picked in the i^{th} subinterval. Thus the height of the i^{th} subinterval would be $f(c_i)$, and the area of the i^{th} rectangle would be $f(c_i)\Delta x_i$.

Summations of rectangles with area $f(c_i)\Delta x_i$ are named after mathematician Georg Friedrich Bernhard Riemann, as given in the following definition.

Definition 1.12: Riemann Sum

Let $f(x)$ be defined on the closed interval $[a,b]$ and let $P = \{x_1, x_2, \ldots, x_{n+1}\}$ be a partition of $[a,b]$, with

$$a = x_1 < x_2 < \ldots < x_n < x_{n+1} = b.$$

Let Δx_i denote the length of the i^{th} subinterval $[x_i, x_{i+1}]$ and let c_i denote any value in the i^{th} subinterval. The sum

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

is a **Riemann sum** *of $f(x)$ on $[a,b]$.*

Riemann sums are typically calculated using one of the three rules we have introduced. The uniformity of construction makes computations easier. Before working another example, let's summarize some of what we have learned in a convenient way.

Riemann Sums Using Rules (Left - Right - Midpoint)

Consider a function $f(x)$ defined on an interval $[a,b]$. The area under this curve is approximated by

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

1. When the n subintervals have equal length, $\Delta x_i = \Delta x = \dfrac{b-a}{n}$.
2. The i^{th} term of the partition is $x_i = a + (i-1)\Delta x$. (This makes $x_{n+1} = b$.)
3. The Left Hand Rule summation is: $\displaystyle\sum_{i=1}^{n} f(x_i)\Delta x$.
4. The Right Hand Rule summation is: $\displaystyle\sum_{i=1}^{n} f(x_{i+1})\Delta x$.
5. The Midpoint Rule summation is: $\displaystyle\sum_{i=1}^{n} f\left(\frac{x_i + x_{i+1}}{2}\right)\Delta x$.

Figure 1.6 shows the approximating rectangles of a Riemann sum. While the rectangles in this example do not approximate well the shaded area, they demonstrate that the subinterval widths may vary and the heights of the rectangles can be determined without following a particular rule.

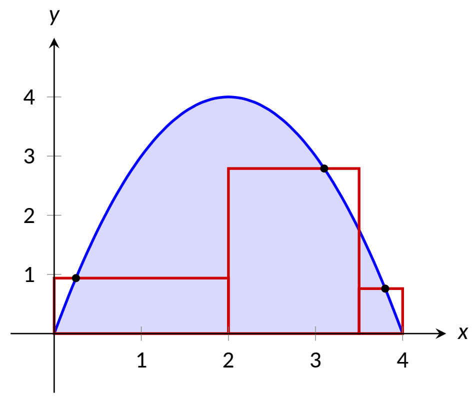


Figure 1.6: General Riemann sum to approximate the area under $f(x) = 4x - x^2$

Let's do another example.

Example 1.13: Approximating Area Using Sums

Approximate the area under $f(x) = (5x+2)$ *on the interval* $[-2,3]$ *using the Midpoint Rule and ten equally spaced intervals.*

Solution. Following the above discussion, we have

$$\Delta x = \frac{3-(-2)}{10} = 1/2$$
$$x_i = (-2) + (1/2)(i-1) = i/2 - 5/2.$$

As we are using the Midpoint Rule, we will also need x_{i+1} and $\dfrac{x_i + x_{i+1}}{2}$. Since $x_i = i/2 - 5/2$,

$$x_{i+1} = (i+1)/2 - 5/2 = i/2 - 2.$$

This gives

$$\frac{x_i + x_{i+1}}{2} = \frac{(i/2 - 5/2) + (i/2 - 2)}{2} = \frac{i - 9/2}{2} = i/2 - 9/4.$$

We now construct the Riemann sum and compute its value using summation formulas.

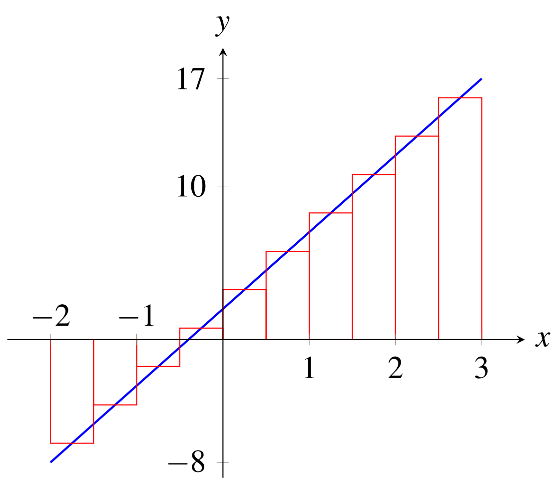
$$\begin{aligned}
\sum_{i=1}^{10} f\left(\frac{x_i + x_{i+1}}{2}\right)\Delta x &= \sum_{i=1}^{10} f(i/2 - 9/4)\Delta x \\
&= \sum_{i=1}^{10} \big(5(i/2 - 9/4) + 2\big)\Delta x \\
&= \Delta x \sum_{i=1}^{10} \left[\left(\frac{5}{2}\right) i - \frac{37}{4}\right]
\end{aligned}$$

$$= \Delta x \left(\frac{5}{2} \sum_{i=1}^{10} (i) - \sum_{i=1}^{10} \left(\frac{37}{4} \right) \right)$$

$$= \frac{1}{2} \left(\frac{5}{2} \cdot \frac{10(11)}{2} - 10 \cdot \frac{37}{4} \right)$$

$$= \frac{45}{2} = 22.5$$

Note the graph below of $f(x) = 5x + 2$ and its area-approximation using the Midpoint Rule and 10 evenly spaced subintervals. The regions whose areas are computed are triangles, meaning we can find the exact answer without summation techniques. We find that the exact answer is indeed 22.5. One of the strengths of the Midpoint Rule is that often each rectangle includes area that should not be counted, but misses other area that should. When the partition width is small, these two amounts are about equal and these errors almost "cancel each other out." In this example, since our function is a line, these errors are exactly equal and they do cancel each other out, giving us the exact answer.



Note too that when the function is negative, the rectangles have a "negative" height and a negative signed area. When we compute the area of the rectangle, we use $f(c_i)\Delta x$; when f is negative, the area is counted as negative. ♣

Notice in the previous example that while we used ten equally spaced intervals, the number "10" didn't play a big role in the calculations until the very end. Mathematicians love abstract ideas; let's approximate the area of another region using n subintervals, where we do not specify a value of n until the very end.

Example 1.14: Approximating Area Using Sums

Revisit $f(x) = 4x - x^2$ on the interval $[0,4]$ yet again. Approximate the area under this curve using the Right Hand Rule with n equally spaced subintervals.

Solution. We know $\Delta x = \frac{4-0}{n} = 4/n$. We also find $x_i = 0 + \Delta x(i-1) = 4(i-1)/n$. The Right Hand Rule uses x_{i+1}, which is $x_{i+1} = 4i/n$. We construct the Right Hand Rule Riemann sum as follows.

$$\sum_{i=1}^{n} f(x_{i+1})\Delta x = \sum_{i=1}^{n} f\left(\frac{4i}{n}\right)\Delta x$$

$$
\begin{aligned}
&= \sum_{i=1}^{n}\left[4\frac{4i}{n} - \left(\frac{4i}{n}\right)^2\right]\Delta x \\
&= \sum_{i=1}^{n}\left(\frac{16\Delta x}{n}\right)i - \sum_{i=1}^{n}\left(\frac{16\Delta x}{n^2}\right)i^2 \\
&= \left(\frac{16\Delta x}{n}\right)\sum_{i=1}^{n} i - \left(\frac{16\Delta x}{n^2}\right)\sum_{i=1}^{n} i^2 \\
&= \left(\frac{16\Delta x}{n}\right)\cdot\frac{n(n+1)}{2} - \left(\frac{16\Delta x}{n^2}\right)\frac{n(n+1)(2n+1)}{6} \qquad \left(\begin{array}{c}\text{recall}\\ \Delta x = 4/n\end{array}\right) \\
&= \frac{32(n+1)}{n} - \frac{32(n+1)(2n+1)}{3n^2} \qquad \text{(now simplify)} \\
&= \frac{32}{3}\left(1 - \frac{1}{n^2}\right)
\end{aligned}
$$

The result is an amazing, easy to use formula. To approximate the area with ten equally spaced subintervals and the Right Hand Rule, set $n = 10$ and compute

$$\frac{32}{3}\left(1 - \frac{1}{10^2}\right) = 10.56.$$

Recall how earlier we approximated the area with 4 subintervals; with $n = 4$, the formula gives 10, our answer as before.

It is now easy to approximate the area with $1{,}000{,}000$ subintervals! Hand-held calculators will round off the answer a bit prematurely giving an answer of 10.66666667. (The actual answer is 10.666666666656.)

We now take an important leap. Up to this point, our mathematics has been limited to geometry and algebra (finding areas and manipulating expressions). Now we apply *calculus*. For any *finite* n, we know that the corresponding Right Hand Rule Riemann sum is:

$$\frac{32}{3}\left(1 - \frac{1}{n^2}\right).$$

Both common sense and high–level mathematics tell us that as n gets large, the approximation gets better. In fact, if we take the *limit* as $n \to \infty$, we get the *exact area*. That is,

$$
\begin{aligned}
\lim_{n\to\infty}\frac{32}{3}\left(1 - \frac{1}{n^2}\right) &= \frac{32}{3}(1-0) \\
&= \frac{32}{3} = 10.\overline{6}
\end{aligned}
$$

This is a fantastic result. By considering n equally–spaced subintervals, we obtained a formula for an approximation of the area that involved our variable n. As n grows large – without bound – the error shrinks to zero and we obtain the exact area. ♣

This section started with a fundamental calculus technique: make an approximation, refine the approximation to make it better, then use limits in the refining process to get an exact answer. That is precisely what we just did.

Let's practice this again.

Example 1.15: Approximating Area With a Formula, Using Sums

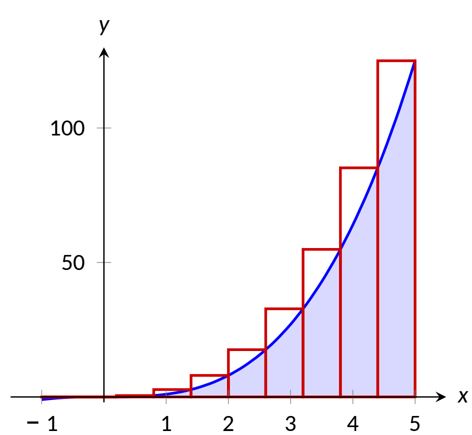
Find a formula that approximates the area under $f(x) = x^3$ on the interval $[-1,5]$ using the Right Hand Rule and n equally spaced subintervals, then take the limit as $n \to \infty$ to find the exact area.

Solution. We have $\Delta x = \frac{5-(-1)}{n} = 6/n$. We have $x_i = (-1) + (i-1)\Delta x$; as the Right Hand Rule uses x_{i+1}, we have $x_{i+1} = (-1) + i\Delta x$.

The Riemann sum corresponding to the Right Hand Rule is (followed by simplifications):

$$\begin{aligned}
\sum_{i=1}^{n} f(x_{i+1})\Delta x &= \sum_{i=1}^{n} f(-1+i\Delta x)\Delta x \\
&= \sum_{i=1}^{n} (-1+i\Delta x)^3 \Delta x \\
&= \sum_{i=1}^{n} \big((i\Delta x)^3 - 3(i\Delta x)^2 + 3i\Delta x - 1\big)\Delta x \\
&= \sum_{i=1}^{n} \big(i^3\Delta x^4 - 3i^2\Delta x^3 + 3i\Delta x^2 - \Delta x\big) \\
&= \Delta x^4 \sum_{i=1}^{n} i^3 - 3\Delta x^3 \sum_{i=1}^{n} i^2 + 3\Delta x^2 \sum_{i=1}^{n} i - \sum_{i=1}^{n} \Delta x \\
&= \Delta x^4 \left(\frac{n(n+1)}{2}\right)^2 - 3\Delta x^3 \frac{n(n+1)(2n+1)}{6} + 3\Delta x^2 \frac{n(n+1)}{2} - n\Delta x \\
&= \frac{1296}{n^4} \cdot \frac{n^2(n+1)^2}{4} - 3\frac{216}{n^3} \cdot \frac{n(n+1)(2n+1)}{6} + 3\frac{36}{n^2}\frac{n(n+1)}{2} - 6 \\
&= 156 + \frac{378}{n} + \frac{216}{n^2}
\end{aligned}$$

Once again, we have found a compact formula for approximating the area with n equally spaced subintervals and the Right Hand Rule. The graph below depicts the graph of f and its area-approximation using the Right Hand Rule and 10 evenly spaced subintervals. This yields an approximation of 195.96. Using $n = 100$ gives an approximation of 159.802.



Now find the exact answer using a limit:

$$\lim_{n\to\infty}\left(156+\frac{378}{n}+\frac{216}{n^2}\right)=156.$$

♣

We have used limits to evaluate exactly given definite limits. Will this always work? We will show, given not–very–restrictive conditions, that yes, it will always work.

The previous two examples demonstrated how an expression such as

$$\sum_{i=1}^{n} f(x_{i+1})\Delta x$$

can be rewritten as an expression explicitly involving n, such as $32/3(1-1/n^2)$.

Viewed in this manner, we can think of the summation as a function of n. An n value is given (where n is a positive integer), and the sum of areas of n equally spaced rectangles is returned, using the Left Hand, Right Hand, or Midpoint Rules.

Given a function $f(x)$ defined on the interval $[a,b]$ let:

- $S_L(n)=\displaystyle\sum_{i=1}^{n} f(x_i)\Delta x$, the sum of equally spaced rectangles formed using the Left Hand Rule,
- $S_R(n)=\displaystyle\sum_{i=1}^{n} f(x_{i+1})\Delta x$, the sum of equally spaced rectangles formed using the Right Hand Rule, and
- $S_M(n)=\displaystyle\sum_{i=1}^{n} f\left(\frac{x_i+x_{i+1}}{2}\right)\Delta x$, the sum of equally spaced rectangles formed using the Midpoint Rule.

The following theorem states that we can use any of our three rules to find the exact value of the area under $f(x)$ on $[a,b]$. It also goes two steps further. The theorem states that the height of each rectangle doesn't have to be determined following a specific rule, but could be $f(c_i)$, where c_i is any point in the i^{th} subinterval, as discussed earlier.

The theorem goes on to state that the rectangles do not need to be of the same width. Using the notation of Definition 1.12, let Δx_i denote the length of the i^{th} subinterval in a partition of $[a,b]$. Now let $||\Delta x||$ represent the length of the largest subinterval in the partition: that is, $||\Delta x||$ is the largest of all the Δx_i's. If $||\Delta x||$ is small, then $[a,b]$ must be partitioned into many subintervals, since all subintervals must have small lengths. "Taking the limit as $||\Delta x||$ goes to zero" implies that the number n of subintervals in the partition is growing to infinity, as the largest subinterval length is becoming arbitrarily small. We then interpret the expression

$$\lim_{||\Delta x||\to 0}\sum_{i=1}^{n} f(c_i)\Delta x_i$$

as "the limit of the sum of rectangles, where the width of each rectangle can be different but getting small, and the height of each rectangle is not necessarily determined by a particular rule." The following theorem states that, for a sufficiently nice function, we can use any of our three rules to find the area under $f(x)$ over $[a,b]$.

Theorem 1.16: Area and the Limit of Riemann Sums

Let $f(x)$ be a continuous function on the closed interval $[a,b]$ and let $S_L(n)$, $S_R(n)$ and $S_M(n)$ be the sums of equally spaced rectangles formed using the Left Hand Rule, Right Hand Rule, and Midpoint Rule, respectively. Then:

1. $\lim_{n\to\infty} S_L(n) = \lim_{n\to\infty} S_R(n) = \lim_{n\to\infty} S_M(n) = \lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\Delta x_i$

2. *The net area under f on the interval $[a,b]$ is equal to* $\lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\Delta x_i$.

3. *The net area under f on the interval $[a,b]$ is equal to* $\lim_{\|\Delta x\|\to 0} \sum_{i=1}^{n} f(c_i)\Delta x_i$.

We summarize what we have learned over the past few sections here.

- Knowing the "area under the curve" can be useful. One common example is: the area under a velocity curve is displacement.
- While we can approximate the area under a curve in many ways, we have focused on using rectangles whose heights can be determined using: the Left Hand Rule, the Right Hand Rule and the Midpoint Rule.
- Sums of rectangles of this type are called Riemann sums.
- The exact value of the area can be computed using the limit of a Riemann sum. We generally use one of the above methods as it makes the algebra simpler.

Exercises for Section 1.3

Exercise 1.3.1 *Find the area under $y = 2x$ between $x = 0$ and any positive value for x.*

Exercise 1.3.2 *Find the area under $y = 4x$ between $x = 0$ and any positive value for x.*

Exercise 1.3.3 *Find the area under $y = 4x$ between $x = 2$ and any positive value for x bigger than 2.*

Exercise 1.3.4 *Find the area under $y = 4x$ between any two positive values for x, say $a < b$.*

Exercise 1.3.5 *Let $f(x) = x^2 + 3x + 2$. Approximate the area under the curve between $x = 0$ and $x = 2$ using 4 rectangles and also using 8 rectangles.*

Exercise 1.3.6 *Let $f(x) = x^2 - 2x + 3$. Approximate the area under the curve between $x = 1$ and $x = 3$ using 4 rectangles.*

1.4 The Definite Integral and FTC

1.4.1. Exploring an Example

We begin by exploring an example. Suppose that an object moves in a straight line so that its speed is $3t$ at time t. How far does the object travel between time $t=a$ and time $t=b$? We don't assume that we know where the object is at time $t=0$ or at any other time. It is certainly true that it is *somewhere,* so let's suppose that at $t=0$ the position is k. Then we know that the position of the object at any time is $3t^2/2+k$. This means that at time $t=a$ the position is $3a^2/2+k$ and at time $t=b$ the position is $3b^2/2+k$. Therefore the change in position is

$$\frac{3b^2}{2}+k-\left(\frac{3a^2}{2}+k\right)=\frac{3b^2}{2}-\frac{3a^2}{2}.$$

Notice that the k drops out; this means that it doesn't matter that we don't know k, it doesn't even matter if we use the wrong k, we get the correct answer.

What about a second approach to this problem? We now want to approximate the change in position between time a and time b. We take the interval of time between a and b, divide it into n subintervals, and approximate the distance traveled during each. The starting time of subinterval number i is now $a+(i-1)(b-a)/n$, which we abbreviate as t_{i-1}, so that $t_0=a$, $t_1=a+(b-a)/n$, and so on. The speed of the object is $f(t)=3t$, and each subinterval is $(b-a)/n=\Delta t$ seconds long. The distance traveled during subinterval number i is approximately $f(t_{i-1})\Delta t$, and the total change in distance is approximately

$$f(t_0)\Delta t+f(t_1)\Delta t+\cdots+f(t_{n-1})\Delta t.$$

The exact change in position is the limit of this sum as n goes to infinity. We abbreviate this sum using sigma notation:

$$\sum_{i=0}^{n-1} f(t_i)\Delta t = f(t_0)\Delta t+f(t_1)\Delta t+\cdots+f(t_{n-1})\Delta t$$

The answer we seek is

$$\lim_{n\to\infty}\sum_{i=0}^{n-1} f(t_i)\Delta t.$$

Since this must be the same as the answer we have already obtained, we know that

$$\lim_{n\to\infty}\sum_{i=0}^{n-1} f(t_i)\Delta t=\frac{3b^2}{2}-\frac{3a^2}{2}.$$

The significance of $3t^2/2$, into which we substitute $t=b$ and $t=a$, is of course that it is a function whose derivative is $f(t)$. As we have discussed, by the time we know that we want to compute

$$\lim_{n\to\infty}\sum_{i=0}^{n-1} f(t_i)\Delta t,$$

it no longer matters what $f(t)$ stands for—it could be a speed, or the height of a curve, or something else entirely. We know that the limit can be computed by finding any function with derivative $f(t)$, substituting

a and b, and subtracting. We summarize this result in a theorem in Section 1.4.3 *The Fundamental Theorem of Calculus*, but first, we introduce the new notation *definite integral* and the terminology associated with it.

1.4.2. Defining the Definite Integral

Definition 1.17: The Definite Integral

Let f be a continuous function defined on the interval $[a,b]$ with partition $P = \{t_0, t_1, \ldots, t_n\}$ of $[a,b]$ and subinterval width Δt. If the limit

$$\lim_{n\to\infty} \sum_{i=0}^{n-1} f(t_i)\,\Delta t$$

exists, then the **definite integral** *of f over $[a,b]$ is defined by*

$$\int_a^b f(t)dt = \lim_{n\to\infty} \sum_{i=0}^{n-1} f(t_i)\,\Delta t$$

where

$\int_a^b f(t)\,dt$ *is read "the integral of f from a to b w.r.t. t",*

the symbol $\int$ is called the **integral sign**,

the value a is the **lower limit of integration**,

the value b is the **upper limit of integration**,

the function f is referred to as the **integrand** *of the integral, and*

the variable t is called the **variable of integration**.

Note:

1. The process of finding the definite integral is called **integration** or **integrating** $f(x)$.
2. If the definite integral of f exists over $[a,b]$, then the function f is **integrable** on $[a,b]$.
3. The definite integral is a *number* and not a function.
4. The value of the definite integral is independent of the variable of integration. Whether the integral is written as
$$\int_a^b f(x)\,dx \quad \text{or} \quad \int_a^b f(t)\,dt \quad \text{or} \quad \int_a^b f(u)\,du,$$
it is still the limit of Riemann sums yielding the same value. Hence, the variable of integration is sometimes referred to as a **dummy variable**.
5. If f is non-negative, then the definite integral represents the **area** of the region under the graph of f on $[a,b]$; otherwise, the definite integral represents the **net area** of the regions under the graph of f on $[a,b]$, which we will summarize in the geometric interpretation below.

We should ask ourselves, when is a function integrable? It turns out that no matter what choices are made in the Riemann sums associated with a continuous function, the Riemann sums always converge to the same limit. This is stated in the following theorem, which we will not prove.

Theorem 1.18: Existence of a Definite Integral

A continuous function on $[a,b]$ is integrable over $[a,b]$.

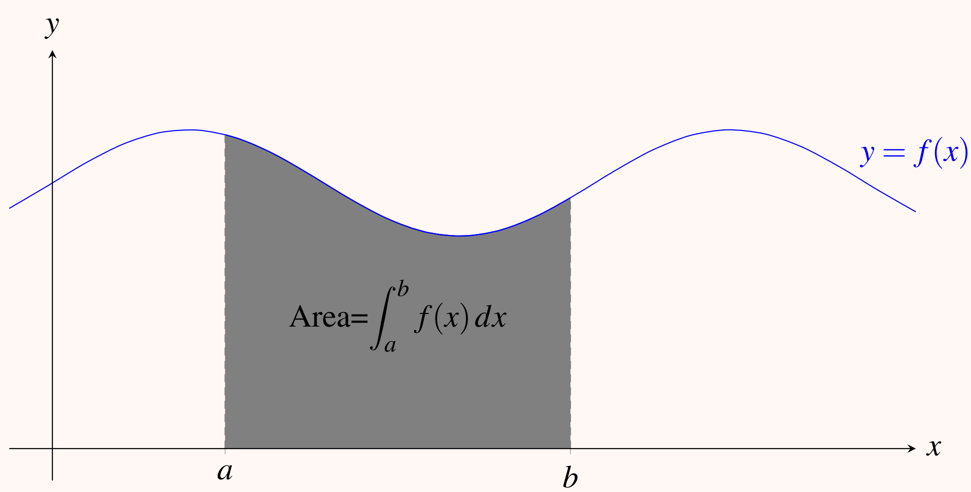
Just like we emphasized the geometric interpretation of the derivative, we do not want to lose sight of the geometric interpretation of the definite integral.

Geometric Interpretation of the Definite Integral – Area

Suppose f is continuous on the interval $[a,b]$, then

$$\int_a^b f(x)\,dx$$

represents the **area** of the region under the graph of f on $[a,b]$ if f is non-negative.

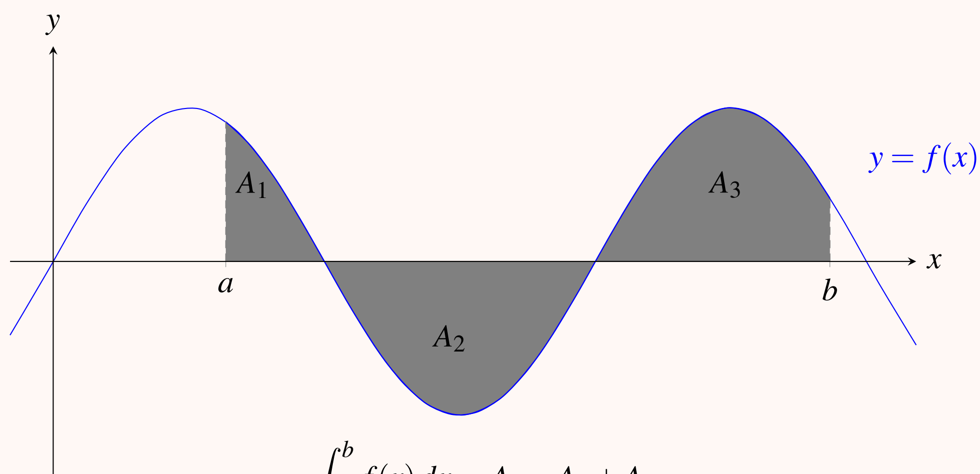


Geometric Interpretation of the Definite Integral – Net Area

Suppose f is continuous on the interval $[a,b]$, then

$$\int_a^b f(x)\,dx$$

represents the **net area** of the region under the graph of f on $[a,b]$.



$$\int_a^b f(x)\,dx = A_1 - A_2 + A_3$$

Lastly, we conclude this section by listing the properties of the definite integral.

Properties of Definite Integrals

Order of limits matters:	$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$
If interval is empty, integral is zero:	$\int_a^a f(x)\,dx = 0$
Constant Multiple Rule:	$\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx$
Sum/Difference Rule:	$\int_a^b [f(x) \pm g(x)]\,dx = \int_a^b f(x)\,dx \pm \int_a^b g(x)\,dx$
Can split up interval $[a,b] = [a,c] \cup [c,b]$:	$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$
The variable does not matter:	$\int_a^b f(x)\,dx = \int_a^b f(t)\,dt$

The reason for the last property is that a definite integral is a *number*, not a function, so the variable is just a placeholder that won't appear in the final answer.

Some additional properties are **comparison** types of properties.

Comparison Properties of Definite Integrals

If $f(x) \geq 0$ for $x \in [a,b]$, then $\int_a^b f(x)\,dx \geq 0.$

If $f(x) \geq g(x)$ for $x \in [a,b]$, then $\int_a^b f(x)\,dx \geq \int_a^b g(x)\,dx.$

If $m \leq f(x) \leq M$ for $x \in [a,b]$, then $m(b-a) \leq \int_a^b f(x)\,dx \leq M(b-a).$

Example 1.19: Properties of Definite Integrals

Suppose $\int_a^b f(x)\ dx = 7$ *and* $\int_a^b g(x)\ dx = 3$. *Find:*

(a) $\int_a^b 2f(x) - 3g(x)\,dx.$

(b) $\int_b^a 2g(x)\,dx.$

(c) $\int_a^a f(x) \cdot g(x)\,dx.$

(d) $\int_a^c f(x)\ dx + \int_c^b f(x)\,dx.$

Solution.

(a) $\int_a^b 2f(x) - 3g(x)\,dx = 2\int_a^b f(x)\,dx - 3\int_a^b g(x)\,dx = 2(7) - 3(3) = 5.$

(b) $\int_b^a 2g(x)\,dx = -2\int_a^b g(x)\,dx = -2(3) = -6.$

(c) $\int_a^a f(x) \cdot g(x)\,dx = 0.$

(d) $\int_a^c f(x)\,dx + \int_c^b f(x)\,dx = \int_a^b f(x)\,dx = 7.$

♣

1.4.3. The Fundamental Theorem of Calculus

We are finally in a position to state the result from our exploration at the beginning of Section 1.4.1. What we have learned is that the integral of f over the interval $[a,b]$ can be computed by finding a function, say $F(t)$, with the property that $F'(t) = f(t)$, and then computing $F(b) - F(a)$. Recall that the function $F(t)$ is called an antiderivative of $f(t)$. Let us now state the theorem:

Theorem 1.20: Fundamental Theorem of Calculus

Suppose that $f(x)$ is continuous on the interval $[a,b]$. If $F(x)$ is any antiderivative of $f(x)$, then

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Let's rewrite this slightly:

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

We've replaced the variable x by t and b by x. These are just different names for quantities, so the substitution doesn't change the meaning of the theorem statement. However, it does allow us to give a new interpretation of the theorem by thinking of the two sides of the equation as functions of x. The expression

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

is a function: plug in a value for x, get out some other value. The expression $F(x) - F(a)$ is of course also a function, and it has a nice property:

$$\frac{d}{dx}(F(x) - F(a)) = F'(x) = f(x),$$

since $F(a)$ is a constant and has derivative zero. In other words, by shifting our point of view slightly, we see that the odd looking function

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

has a derivative, and that in fact $G'(x) = f(x)$.

Note: This is really just a restatement of the Fundamental Theorem of Calculus, and indeed is often called the Fundamental Theorem of Calculus. To avoid confusion, some people call the two versions of the theorem "The Fundamental Theorem of Calculus, part I" and "The Fundamental Theorem of Calculus, part II", although unfortunately there is no universal agreement as to which is part I and which part II. Since it really is the same theorem, differently stated, some people simply call them both "The Fundamental Theorem of Calculus."

Theorem 1.21: Fundamental Theorem of Calculus

Suppose that $f(x)$ is continuous on the interval $[a,b]$ and let

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

Then $G'(x) = f(x)$.

We have not really proved the Fundamental Theorem. In a nutshell, we gave the following argument to justify it: Suppose we want to know the value of

$$\int_a^b f(t)\,dt = \lim_{n\to\infty} \sum_{i=0}^{n-1} f(t_i)\,\Delta t.$$

We can interpret the right hand side as the distance traveled by an object whose speed is given by $f(t)$. We know another way to compute the answer to such a problem: find the position of the object by finding an antiderivative of $f(t)$, then substitute $t = a$ and $t = b$ and subtract to find the distance traveled. This must be the answer to the original problem as well, even if $f(t)$ does not represent a speed.

What's wrong with this? In some sense, nothing. As a practical matter it is a very convincing argument, because our understanding of the relationship between speed and distance seems to be quite solid. From the point of view of mathematics, however, it is unsatisfactory to justify a purely mathematical relationship by appealing to our understanding of the physical universe, which could, however unlikely it is in this case, be wrong.

A complete proof is a bit too involved to include here, but we will indicate how it goes. First, if we can prove the second version of the Fundamental Theorem of Calculus, Theorem 1.21, then we can prove the first version from that:

Proof. Proof of Theorem 1.20.

We know from Theorem 1.21 that

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

is an antiderivative of $f(x)$, and therefore any antiderivative $F(x)$ of $f(x)$ is of the form $F(x) = G(x) + k$. Then

$$\begin{aligned} F(b) - F(a) = G(b) + k - (G(a) + k) &= G(b) - G(a) \\ &= \int_a^b f(t)\,dt - \int_a^a f(t)\,dt. \end{aligned}$$

It is clear that $\int_a^a f(t)\,dt = 0$, so this means that

$$F(b) - F(a) = \int_a^b f(t)\,dt,$$

which is exactly what Theorem 1.20 says. ♣

So the real job is to prove Theorem 1.21. We will sketch the proof, using some facts that we do not prove. First, the following identity is true of integrals:

$$\int_a^b f(t)\,dt = \int_a^c f(t)\,dt + \int_c^b f(t)\,dt$$

This can be proved directly from the definition of the integral, that is, using the limits of sums. It is quite easy to see that it must be true by thinking of either of the two applications of integrals that we have seen. It turns out that the identity is true no matter what c is, but it is easiest to think about the meaning when $a \le c \le b$.

First, if $f(t)$ represents a speed, then we know that the three integrals represent the distance traveled between time a and time b; the distance traveled between time a and time c; and the distance traveled between time c and time b. Clearly the sum of the latter two is equal to the first of these.

Second, if $f(t)$ represents the height of a curve, the three integrals represent the area under the curve between a and b; the area under the curve between a and c; and the area under the curve between c and b. Again it is clear from the geometry that the first is equal to the sum of the second and third.

Proof. Proof of Theorem 1.21.

We want to compute $G'(x)$, so we start with the definition of the derivative in terms of a limit:

$$\begin{aligned}
G'(x) &= \lim_{\Delta x\to 0} \frac{G(x+\Delta x)-G(x)}{\Delta x} \\
&= \lim_{\Delta x\to 0} \frac{1}{\Delta x}\left(\int_a^{x+\Delta x} f(t)\,dt - \int_a^x f(t)\,dt\right) \\
&= \lim_{\Delta x\to 0} \frac{1}{\Delta x}\left(\int_a^{x} f(t)\,dt + \int_x^{x+\Delta x} f(t)\,dt - \int_a^x f(t)\,dt\right) \\
&= \lim_{\Delta x\to 0} \frac{1}{\Delta x}\int_x^{x+\Delta x} f(t)\,dt.
\end{aligned}$$

Now we need to know something about

$$\int_x^{x+\Delta x} f(t)\,dt$$

when Δx is small; in fact, it is very close to $\Delta x f(x)$, but we will not prove this. Once again, it is easy to believe this is true by thinking of our two applications: The integral

$$\int_x^{x+\Delta x} f(t)\,dt$$

can be interpreted as the distance traveled by an object over a very short interval of time. Over a sufficiently short period of time, the speed of the object will not change very much, so the distance traveled will be approximately the length of time multiplied by the speed at the beginning of the interval, namely, $\Delta x f(x)$. Alternately, the integral may be interpreted as the area under the curve between x and $x+\Delta x$. When Δx is very small, this will be very close to the area of the rectangle with base Δx and height $f(x)$; again this is $\Delta x f(x)$. If we accept this, we may proceed:

$$\lim_{\Delta x\to 0} \frac{1}{\Delta x}\int_x^{x+\Delta x} f(t)\,dt = \lim_{\Delta x\to 0} \frac{\Delta x f(x)}{\Delta x} = f(x),$$

which is what we wanted to show. ♣

It is still true that we are depending on an interpretation of the integral to justify the argument, but we have isolated this part of the argument into two facts that are not too hard to prove. Once the last reference to interpretation has been removed from the proofs of these facts, we will have a real proof of the Fundamental Theorem of Calculus.

Note: Now we know that to solve certain kinds of problems, those that lead to a sum of a certain form, we "merely" find an antiderivative and substitute two values and subtract. Unfortunately, finding antiderivatives can be quite difficult. While there are a small number of rules that allow us to compute the derivative of any common function, there are no such rules for antiderivatives. There are some techniques that frequently prove useful as you will see in Chapter 2 *Techniques of Integration*, but we will never be able to reduce the problem to a completely mechanical process.

1.4.4. Notation when Computing a Definite Integral

When we compute a definite integral, we first find an antiderivative of the integrand and then substitute. It is convenient to first display the antiderivative and then do the substitution; we need a notation indicating that the substitution is yet to be done. A typical solution would look like this:

$$\int_1^2 x^2\,dx = \left.\frac{x^3}{3}\right|_1^2 = \frac{2^3}{3} - \frac{1^3}{3} = \frac{7}{3}.$$

The vertical line with subscript and superscript is used to indicate the operation "substitute and subtract" that is needed to finish the evaluation.

Common Mistakes:

1. Dropping the dx at the end of the integral. This is required! Think of the integral as a set of parenthesis consisting of the integral sign and the dx. Both are required so it is clear where the integrand ends and what variable you are integrating with respect to.

$$\int_a^b f(x)\,dx \neq \int_a^b f(x)$$

2. Checking the antiderivative for the integrand using differentiation before following through on the lower and upper bounds of integration.

$$\int_a^b f(x)\,dx = F(x)\Big|_a^b \quad \text{with} \quad F'(x) = f(x)$$

3. Dropping the lower and upper bounds of integration during the solution process. For example:

$$\int_1^2 x^2\,dx \neq \int x^2 dx \neq \left.\frac{x^3}{3}\right|_1^2 = \frac{2^3}{3} - \frac{1^3}{3} = \frac{7}{3}$$

4. Forgetting to evaluate using the lower and upper bounds of integration that lead to the value of the definite integral and instead producing a function. See the note immediately after the definition of the definite integral. For example:

$$\int_1^2 x^2\,dx \neq \frac{x^3}{3}$$

5. Switching the order in which the lower and upper bounds of integration are being dealt with. For example:

$$\int_1^2 x^2\,dx = \left.\frac{x}{3}\right|_1^2 \neq \frac{1^3}{3} - \frac{2^3}{3} = -\frac{7}{3}, \text{ or}$$

$$\int_1^2 x^2\,dx \neq \left.\frac{x}{3}\right|_2^1 = \frac{1^3}{3} - \frac{2^3}{3} = -\frac{7}{3}.$$

1.4.5. Computing a Definite Integral

We seem to have found a pattern when dealing with power functions. When attempting to solve a previous question, we found the antiderivative of x^2 to be $x^3/3+c$ (as it was when solving the indefinite integral). Likewise, when we first began, we were trying to determine a position based on velocity, and $3t$ gave rise to $3t^2/2+k$.

As will be formalized later, we see that in these cases, the power is increased to $n+1$, but we also divide through by this factor, $n+1$. So the antiderivative of x becomes $x^2/2$, the antiderivative of x^2 becomes $x^3/3$, and the antiderivative of x^3 will become $x^4/4$.

Now we will also try with negative and fractional values in the following example.

Example 1.22: Fundamental Theorem of Calculus

Evaluate $\displaystyle\int_1^4 \left(x^3+\sqrt{x}+\frac{1}{x^2}\right)\,dx.$

Solution.

$$\begin{aligned}
\int_1^4 \left(x^3+\sqrt{x}+\frac{1}{x^2}\right)dx &= \left.\frac{x^4}{4}+\frac{2x^{3/2}}{3}-x^{-1}\right|_1^4 \\
&= \left(\frac{(4)^4}{4}+\frac{2(4)^{3/2}}{3}-4^{-1}\right) \\
&\quad -\left(\frac{(1)^4}{4}+\frac{2(1)^{3/2}}{3}-1^{-1}\right) \\
&= \frac{415}{6}
\end{aligned}$$

♣

Note: The above integral used parentheses around the integrand for clarity, but we can leave them out once we are familiar with the notation as shown in the next example.

We next evaluate a definite integral using three different techniques.

Example 1.23: Three Different Techniques

Evaluate $\displaystyle\int_0^2 x+1\,dx$ *by*

(a) Using FTC II (the shortcut)

(b) Using the definition of a definite integral (the limit sum definition)

(c) Interpreting the problem in terms of areas (graphically)

Solution.

(a) The shortcut (FTC II) is the method of choice as it is the fastest. Integrating and using the '*top minus bottom*' rule we have:

$$\begin{aligned}\int_0^2 x+1\,dx &= \left.\frac{x^2}{2}+x\right|_0^2\\ &= \left[\frac{2^2}{2}+2\right]-\left[\frac{0^2}{2}+0\right]=4.\end{aligned}$$

(b) We now use the definition of a definite integral. We divide the interval $[0,2]$ into n subintervals of equal width Δx, and from each interval choose a point x_i^*. Using the formulas

$$\Delta x=\frac{b-a}{n} \qquad \text{and} \qquad x_i=a+i\Delta x,$$

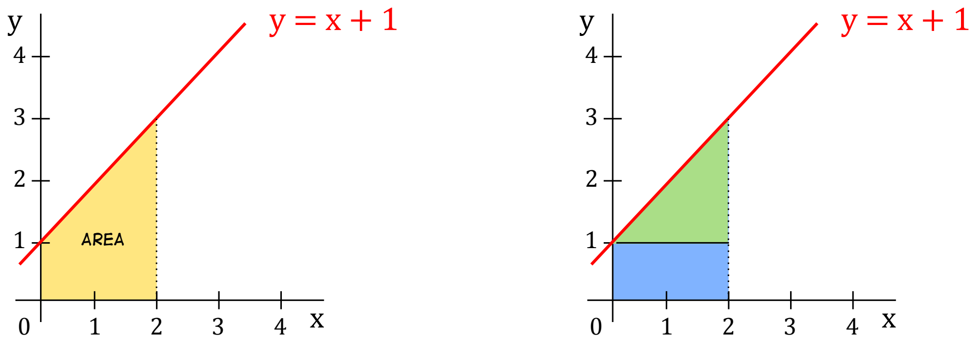
we have

$$\Delta x=\frac{2}{n} \qquad \text{and} \qquad x_i=0+i\Delta x=\frac{2i}{n}.$$

Then taking x_i^*'s as right endpoints for convenience (so that $x_i^*=x_i$), we have:

$$\begin{aligned}\int_0^2 x+1\,dx &= \lim_{n\to\infty}\sum_{i=1}^n f(x_i^*)\,\Delta x\\ &= \lim_{n\to\infty}\sum_{i=1}^n f\left(\frac{2i}{n}\right)\frac{2}{n}\\ &= \lim_{n\to\infty}\sum_{i=1}^n \left(\frac{2i}{n}+1\right)\frac{2}{n}\\ &= \lim_{n\to\infty}\sum_{i=1}^n \left(\frac{4i}{n^2}+\frac{2}{n}\right)\\ &= \lim_{n\to\infty}\left(\sum_{i=1}^n \frac{4i}{n^2}+\sum_{i=1}^n\frac{2}{n}\right)\\ &= \lim_{n\to\infty}\left(\frac{4}{n^2}\sum_{i=1}^n i+\frac{2}{n}\sum_{i=1}^n 1\right)\\ &= \lim_{n\to\infty}\left(\frac{4}{n^2}\frac{n(n+1)}{2}+\frac{2}{n}n\right)\\ &= \lim_{n\to\infty}\left(2+\frac{2}{n}+2\right)\\ &= 4.\end{aligned}$$

(c) Finally, let's evaluate the net area under $x+1$ from 0 to 2.



Thus, the area is the sum of the areas of a rectangle and a triangle. Hence,

$$\begin{aligned}\int_0^2 x+1\,dx &= \text{Net Area}\\ &= \text{Area of rectangle} + \text{Area of triangle}\\ &= (2)(1)+\frac{1}{2}(2)(2)\\ &= 4.\end{aligned}$$

♣

Example 1.24: FTC and Marginal Profit

A small bakery determines that the marginal profit function for selling q loaves of bread is

$$P'(q) = -0.0005q^2 + 2.5$$

dollars per day per loaf. If the fixed production cost is \$10,

(a) determine $P(50)$, and

(b) determine whether or not it is advisable to increase production from 50 loaves to 100 loaves.

Solution.

(a) We first notice that

$$\int_a^{50} P'(q)\,dq = P(50) - P(a).$$

Since we know that $P(0) = -10$ (the fixed cost), we take $a = 0$ and integrate:

$$\begin{aligned}\int_0^{50} P'(q)\,dq &= \int_0^{50} \left(-0.0005q^2 + 2.5\right)\,dq \\ &= -0.0005 \int_0^{50} q^2 dq + 2.5 \int_0^{50} dq \\ &= -0.0005 \frac{q^3}{3}\bigg|_0^{50} + 2.5q\bigg|_0^{50} \\ &= -0.0005\left(\frac{50^3}{3} - 0\right) + 2.5\,(50-0) = 104.167\end{aligned}$$

Thus,

$$P(50) = 104.167 - (-10) = 114.167.$$

Therefore, the daily profit the bakery realizes from selling 50 loaves of bread is approximately \$114.

(b) The change in the profit can be determined by integrating

$$\begin{aligned}\int_{50}^{100} P'(q)\,dq &= \int_{50}^{100} \left(-0.0005q^2 + 2.5\right)\,dq \\ &= -0.0005\left(\frac{100^3}{3} - \frac{50^3}{3}\right) + 2.5\,(100-50) = -20.8333\end{aligned}$$

This means that the bakery would lose about \$21 per day if production were doubled. Therefore it is not advisable to increase production to 100 loaves.

♣

We next apply FTC to differentiate a function.

Example 1.25: Using FTC

Differentiate the following function:

$$g(x) = \int_{-2}^{x} \cos(1+5t)\sin t\,dt$$

Solution. We simply apply the Fundamental Theorem of Calculus directly to get:

$$g'(x) = \cos(1+5x)\sin x.$$

♣

Using the Chain Rule we can derive a formula for some more complicated problems when the upper limit of integration is some function of x rather than simply x. We have

$$\frac{d}{dx}\int_a^{v(x)} f(t)\,dt = f(v(x)) \cdot v'(x).$$

Now what if the upper limit is constant and the lower limit is a function of x? Then we interchange the limits and add a minus sign to get

$$\frac{d}{dx}\int_{u(x)}^{a} f(t)\,dt = -\frac{d}{dx}\int_{a}^{u(x)} f(t)\,dt = -f(u(x))\cdot u'(x).$$

Combining these two we can get a formula where both limits are functions of x. We break up the integral as follows:

$$\int_{u(x)}^{v(x)} f(t)\,dt = \int_{u(x)}^{a} f(t)\,dt + \int_{a}^{v(x)} f(t)\,dt$$

We just need to make sure $f(a)$ exists after we break up the integral. Then differentiating and using the above two formulas gives:

FTC and Chain Rule Formula:

$$\frac{d}{dx}\int_{u(x)}^{v(x)} f(t)\,dt = f(v(x))v'(x) - f(u(x))u'(x)$$

Many textbooks do not show this formula and instead to solve these types of problems will use FTC I along with the tricks we used to derive the formula above. Either method is perfectly fine.

Example 1.26: FTC and Chain Rule I

Differentiate the following integral:

$$\int_{10x}^{x^2} t^3\sin(1+t)\,dt$$

Solution. We will use the formula above. We have $f(t) = t^3\sin(1+t)$, $u(x) = 10x$ and $v(x) = x^2$. Then $u'(x) = 10$ and $v'(x) = 2x$. Thus,

$$\begin{aligned}\frac{d}{dx}\int_{10x}^{x^2} t^3\sin(1+t)\,dt &= (x^2)^3\sin(1+(x^2))(2x) - (10x)^3\sin(1+(10x))(10) \\ &= 2x^7\sin(1+x^2) - 10000x^3\sin(1+10x)\end{aligned}$$

♣

Example 1.27: FTC and Chain Rule II

Differentiate the following integral with respect to x:

$$\int_{x^3}^{2x} 1+\cos t\,dt$$

Solution. Using the formula we have:

$$\frac{d}{dx}\int_{x^3}^{2x} 1+\cos t\,dt = (1+\cos(2x))(2)-(1+\cos(x^3))(3x^2).$$

♣

Exercises for Section 1.4

Exercise 1.4.1 *Determine the area under the graph of each function on the given interval.*

(a) $f(x)=3x+1$ on $[0,2]$

(b) $f(t)=2t$ on $[1,e]$

(c) $g(t)=1+\sqrt{t}$ on $[3,6]$

(d) $h(x)=\sin(x)$ on $[0,\pi]$

Exercise 1.4.2 *Evaluate the following definite integrals.*

(a) $\displaystyle\int_1^4 t^2+3t\,dt$

(b) $\displaystyle\int_0^\pi \sin t\,dt$

(c) $\displaystyle\int_1^{10}\frac{1}{x}\,dx$

(d) $\displaystyle\int_0^5 e^x\,dx$

(e) $\displaystyle\int_0^3 x^3\,dx$

(f) $\displaystyle\int_1^2 x^5\,dx$

(g) $\displaystyle\int_1^2 7x^{5/2}\,dx$

(h) $\displaystyle\int_0^2 \left(2x^3+x^2-5\right)\,dx$

(i) $\displaystyle\int_0^2\left(\frac{1}{\sqrt{x}}-\sqrt{x}\right)\,dx$

(j) $\displaystyle\int_2^5\frac{1}{3x+2}\,dx$

(k) $\displaystyle\int_{-\pi/3}^0\frac{\tan x}{\cos x}\,dx$

Exercise 1.4.3 *Find the derivative of the given function G.*

(a) $G(x)=\displaystyle\int_1^x t^2-3t\,dt$

(b) $G(x)=\displaystyle\int_1^{x^2} t^2-3t\,dt$

(c) $G(x)=\displaystyle\int_1^x e^{t^2}\,dt$

(d) $G(x)=\displaystyle\int_1^{x^2} e^{t^2}\,dt$

(e) $G(x)=\displaystyle\int_1^x \tan(t^2)\,dt$

(f) $G(x)=\displaystyle\int_{10x}^{x^2} \tan(t^2)\,dt$

Exercise 1.4.4 *Suppose* $\displaystyle\int_1^4 f(x)\,dx=2$ *and* $\displaystyle\int_1^4 g(x)\,dx=7$. *Find* $\displaystyle\int_1^4 (5f(x)+3g(x))\,dx$ *and* $\displaystyle\int_1^4 (6-2f(x))\,dx$.

Exercise 1.4.5 *Suppose* $\displaystyle\int_{-2}^5 f(x)\,dx=3$ *and* $\displaystyle\int_1^5 f(x)\,dx=-2$. *Find* $\displaystyle\int_{-2}^1 f(x)\,dx$.

Exercise 1.4.6 *Suppose* $\int_{-1}^{1} f(x)\,dx = 10$ *and* $\int_{0}^{1} f(x)\,dx = 2$. *Find* $\int_{-1}^{0} 2f(x)\,dx$.

Exercise 1.4.7 *A manufacturer determines that the marginal cost function associated with the production of q units is*

$$C'(q) = 0.0006q^2 - 0.5q + 25$$

dollars per unit per day. Suppose the daily fixed cost is \$100.

(a) Determine the total cost of producing 30 units.

(b) After 30 units have been made, determine the cost of producing an additional 20 units.

Exercise 1.4.8 *A manufacturer determines that the marginal revenue function associated with the production of q units is*

$$R'(q) = -0.5q + 100$$

dollars per unit per day.

(a) Determine the total revenue from selling 100 units.

(b) After 100 units have been made, determine the revenue from selling an additional 100 units.

1.5 Indefinite Integrals

In this section we focus on the indefinite integral: its definition, the differences between the definite and indefinite integrals, some basic integral rules, and how to compute a definite integral.

1.5.1. Defining the Indefinite Integral

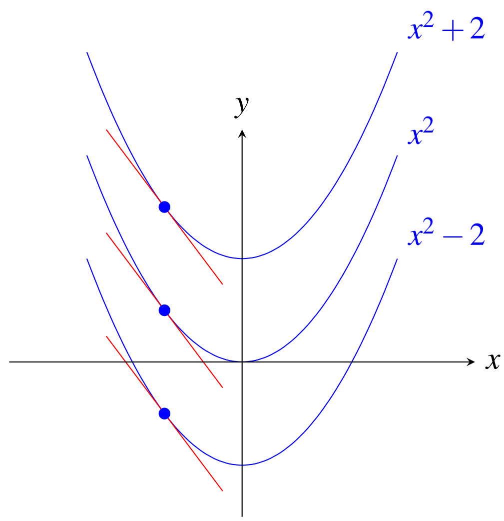
Recall the definition of the antiderivative from Section 1.1: A function F is an **antiderivative** of f on an interval I if $F'(x) = f(x)$ for all x in I. Let us explore the antiderivative concretely by letting $f(x) = 2x$. Then we can readily determine that the antiderivative of f is the function $F(x) = x^2$, i.e. $F'(x) = f(x)$. However, the function $F(x) + 1 = x^2 + 1$ also has f as its derivative:

$$\frac{d}{dx}(F(x) + 1) = \frac{d}{dx}\left(x^2 + 1\right) = 2x.$$

In fact, any function $F(x) + C = x^2 + C$ for some real constant C has f as its derivative:

$$\frac{d}{dx}(F(x) + C) = \frac{d}{dx}\left(x^2 + C\right) = 2x.$$

It comes as no surprise to us that the graphs of the family of functions $F(x) + C$ are visually just vertical displacements of $F(x)$. In the particular case when $F(x) = x^2$, we can also see that the graphs of the family of functions $F(x) + C$ below that at any point x the tangent lines are parallel, i.e. the tangent slopes are the same, i.e. the family of functions has the same derivative $f(x) = 2x$:



This leads us to the following result:

Definition 1.28: General Antiderivative

If a function F is an **antiderivative** *of f on an interval I, then the most* **general antiderivative** *of f on an interval I is*

$$F(x)+C$$

where C is any real constant.

Let us now define the **indefinite integral**.

Definition 1.29: The Indefinite Integral

The set of all antiderivatives of a function $f(x)$ is the **indefinite integral** *of $f(x)$ with respect to x and denoted by*

$$\int f(x)\,dx,$$

where

$\int f(x)\,dx$ is read "the integral of f w.r.t. x",

the symbol $\int$ is called the **integral sign**,

the function f is referred to as the **integrand** *of the integral, and*

the variable x is called the **variable of integration**.

Note:

1. The process of finding the indefinite integral is also called **integration** or **integrating** $f(x)$.
2. The above definition says that if a function F is an antiderivative of f, then

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

for some real constant C.

3. Unlike the definite integral, the indefinite integral is a *function*.

1.5.2. Definite Integral versus Indefinite Integral

Due to the close relationship between an integral and an antiderivative, the integral sign is also used to mean "antiderivative". You can tell which is intended by whether the limits of integration are included:

$$\int_1^2 x^2\,dx$$

is an ordinary integral, also called a definite integral, because it has a definite value, namely

$$\int_1^2 x^2\,dx = \frac{2^3}{3} - \frac{1^3}{3} = \frac{7}{3}.$$

We use

$$\int x^2\,dx$$

to denote the antiderivative of x^2, also called an indefinite integral. So this is evaluated as

$$\int x^2\,dx = \frac{x^3}{3} + C,$$

which is clearly a function as opposed to the definite integral which is a value. It is customary to include the constant C to indicate that there are really an infinite number of antiderivatives. We do not need this C to compute definite integrals, but in other circumstances we will need to remember that the C is there, so it is best to get into the habit of writing the C.

1.5.3. Computing Indefinite Integrals

We are finally ready to compute some indefinite integrals and introduce some basic integration rules from our knowledge of derivatives. We will first point out some common mistakes frequently observed in student work.

Common Mistakes:

1. Dropping the dx at the end of the integral. This is required! Think of the integral as a set of parenthesis consisting of the integral sign and the dx. Both are required so it is clear where the integrand ends and what variable you are integrating with respect to.

$$\int f(x)\,dx \neq \int f(x)$$

2. Forgetting the $+C$ during the solution process and thereby not showing that the solution of an indefinite integration process is the set of all antiderivatives of the integrand. As an aside, constants of integration play a major role in the topic of *Differential Equations*. For example:

$$\int x^2 dx \neq \frac{x^3}{3}$$

Caution: Note that we don't have properties to deal with products or quotients of functions, that is,

$$\int f(x) \cdot g(x)\, dx \neq \int f(x)\, dx \int g(x)\, dx.$$

$$\int \frac{f(x)}{g(x)}\, dx \neq \frac{\int f(x)\, dx}{\int g(x)\, dx}.$$

With derivatives, we had the product and quotient rules to deal with these cases. For integrals, we have no such rules, but we will learn a variety of different techniques to deal with these cases.

The following integral rules can be proved by taking the derivative of the functions on the right side.

Integral Rules

Constant Rule: $\int k\, dx = kx + C.$

Constant Multiple Rule: $\int k f(x)\, dx = k \int f(x)\, dx$, where k constant.

Sum/Difference Rule: $\int f(x) \pm g(x)\, dx = \int f(x)\, dx \pm \int g(x)\, dx.$

Power Rule: $\int x^n\, dx = \frac{x^{n+1}}{n+1} + C, \quad n \neq -1.$

Log Rule: $\int \frac{1}{x}\, dx = \ln|x| + C, \quad x \neq 0.$

Exponent Rule: $\int a^{kx} = \frac{a^{kx}}{k \ln a} + C, \quad x \neq 0.$

Trig Rules:

$$\int \sin x\, dx = -\cos x + C,$$

$$\int \cos x\, dx = \sin x + C,$$

$$\int \sec^2 x\, dx = \tan x + C,$$

$$\int \sec x \tan x\, dx = \sec x + C,$$

$$\int \csc^2 x\, dx = -\cot x + C,$$

$$\int \csc x \cot x\, dx = \csc x + C,$$

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

$$\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C.$$

Example 1.30: Indefinite Integral

If $f'(x) = x^4 + 2x - 8\sin x$ then what is $f(x)$?

Solution. The answer is:

$$\begin{aligned} f(x)=\int f'(x)\,dx &= \int \left(x^4+2x-8\sin x\right)\,dx \\ &= \int x^4\,dx+2\int x\,dx-8\int \sin x\,dx \\ &= \frac{x^5}{5}+x^2+8\cos x+C, \end{aligned}$$

where C is a constant. ♣

Example 1.31: Indefinite Integral

Find the indefinite integral $\int 3x^2\,dx$.

Solution.

$$\begin{aligned} \int 3x^2\,dx &= 3\int x^2\,dx \\ &= 3\frac{x^3}{3}+C \\ &= x^3+C \end{aligned}$$

♣

Example 1.32: Indefinite Integral

Find the indefinite integral $\int \frac{2}{\sqrt{x}}\,dx$.

Solution.

$$\begin{aligned} \int \frac{2}{\sqrt{x}}\,dx &= 2\int x^{-\frac{1}{2}}\,dx \\ &= 2\frac{x^{-\frac{1}{2}+1}}{-\frac{1}{2}+1}+C \\ &= 4\sqrt{x}+C \end{aligned}$$

♣

Note: There is an implicit (built-in) restriction in the above calculations, namely $x \geq 0$. This comes from the fact that the integrand $\frac{2}{\sqrt{x}}$ is only defined for $x \geq 0$, and, not surprisingly, the indefinite integral is also only defined over this interval.

Example 1.33: Indefinite Integral

Find the indefinite integral $\int \left(\frac{1}{x}+e^{7x}+x^{\pi}+7\right)\,dx.$

Solution.

$$\begin{aligned}\int \left(\frac{1}{x}+e^{7x}+x^{\pi}+7\right)\,dx &= \int \frac{1}{x}\,dx+\int e^{7x}\,dx+\int x^{\pi}\,dx+\int 7\,dx \\ &= \ln|x|+\frac{1}{7}e^{7x}+\frac{x^{\pi+1}}{\pi+1}+7x+C\end{aligned}$$

♣

Note: Just like in the previous example, the rule for integration $\frac{1}{x}$ also comes with an implicit restriction, namely $x \neq 0$. The function $\frac{1}{x}$ clearly has a well defined area, as long as you stick to values of $x<0$ or $x>0$ and is undefined for $x=0$, but $\ln x$ is undefined for $x \leq 0$. However, we know from differential calculus that $\frac{d}{dx}\ln|x| = \frac{1}{x}$ (try this derivative again to convince yourself by rewriting this function as a peicewise-defined function). Hence, the antiderivative of $\frac{1}{x}$ becomes $\ln|x|$.

Example 1.34: Finding Cost Functions

Suppose a publishing company has found that the marginal cost at a level of production of x thousand magazines is given by

$$C'(x) = \frac{25}{\sqrt{x}}$$

and that the fixed cost, i.e. the cost before the first book can be produced, is $36,000. Find the cost function $C(x)$.

Solution. By the indefinite integral rules

$$\int \frac{25}{\sqrt{x}}dx = \int 25x^{-1/2}dx = 25(2x^{1/2})+k = 50x^{1/2}+k,$$

where k represents the constant of integration to avoid confusion with the cost function. Notice that the production x is always non-negative, and so we proceed with the integration with the implicit assumption that $x \geq 0$ is automatically satisfied.

To find the value of k, use the fact that $C(0)$ is $36,000$.

$$\begin{aligned}C(x) &= 50x^{1/2}+k \\ 36,000 &= 50\cdot 0+k \\ k &= 36,000\end{aligned}$$

With this result, the cost function is $C(x) = 50x^{1/2}+36,000$. ♣

Example 1.35: Finding Revenue and Demand Functions

Suppose the marginal revenue from a product is given by

$$800e^{-0.2}+7.5.$$

(a) Find the revenue function for this product.

(b) Find the demand function for this product.

Solution.

(a) The marginal revenue is the derivative of the revenue function, so

$$\begin{aligned} R'(q) &= 800e^{-0.2q}+7.5 \\ R(q) &= \int \left(800e^{-0.2q}+7.5\right)dq \\ &= 800\frac{e^{-0.2q}}{-0.2}+7.5q+C \\ &= -4000e^{-0.2q}+7.5q+C \end{aligned}$$

If no items are sold, then there is no revenue. Hence, $q=0$ and $R=0$, and so

$$\begin{aligned} 0 &= -4000e^{-0.2(0)}+7.5(0)+C \\ 0 &= -4000+C \\ C &= 4000 \end{aligned}$$

Therefore, the revenue function is

$$R(q)=-4000e^{-0.2q}+7.5q+4000.$$

(b) Recall that $R=qp$, where p is the demand function that represents the price p as a function of q. So

$$\begin{aligned} -4000e^{-0.2q}+7.5q+4000 &= qp \\ \frac{-4000e^{-0.2q}+7.5q+4000}{q} &= p \end{aligned}$$

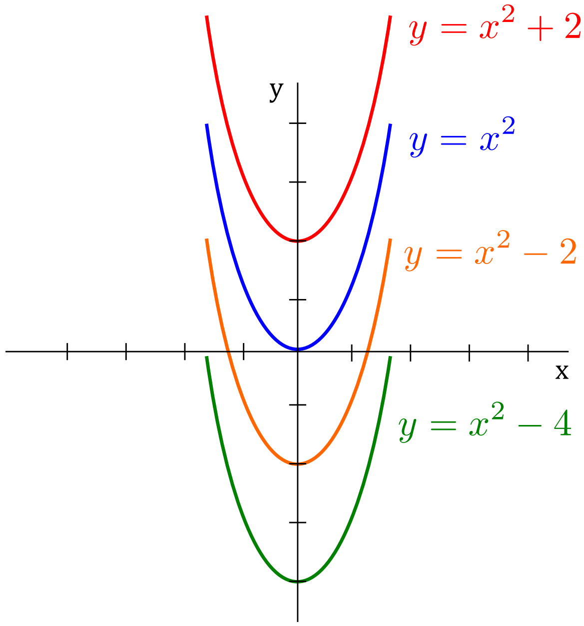
Therefore, the demand function is

$$p(q)=\frac{-4000e^{-0.2q}+7.5q+4000}{q}.$$

♣

1.5.4. Differential Equations and Constants of Integration

An equation involving derivatives where we want to solve for the original function is called a **differential equation**. For example, $f'(x) = 2x$ is a differential equation with general solution $f(x) = x^2 + C$. Some solutions (i.e., specific values of C) are shown below.



As seen with integral curves, we may have an infinite family of solutions satisfying the differential equation. However, if we were given a point (called an *initial value*) on the curve then we could determine $f(x)$ completely. Such a problem is known as an *initial value problem*.

Example 1.36: Initial Value Problem

If $f'(x) = 2x$ and $f(0) = 2$ then determine $f(x)$.

Solution. As previously stated, we have a solution of:

$$f(x) = x^2 + C.$$

But $f(0) = 2$ implies:

$$2 = 0^2 + C \quad \rightarrow \quad C = 2.$$

Therefore, $f(x) = x^2 + 2$ is the solution to the initial value problem. ♣

Exercises for Section 1.5

Exercise 1.5.1 *Find the following indefinite integrals.*

(a) $\int 8\sqrt{x}dx$

(b) $\int 3t^2+1dt$

(c) $\int 4/\sqrt{y}dy$

(d) $\int 2/z^2dz$

(e) $\int 7s^{-1}ds$

(f) $\int (5x+1)^2dx$

(g) $\int (t-6)^2dt$

(h) $\int z^{3/2}dz$

(i) $\int \frac{2}{x\sqrt{x}}dx$

(j) $\int |2t-4|dt$

(k) $\int (\cos(2x)+4\sin(x))\,dx$

(l) $\int \left(2y-\sec^2 y\right)dy$

(m) $\int \left(\frac{\tan x}{\cos x}-\csc^2 x\right)dx$

Exercise 1.5.2 *Find the following indefinite integrals.* **Hint:** *Simplify the integrand first.*

(a) $\int \frac{2s^3+s^2-s}{s}ds$

(b) $\int \frac{3x^3-x^2+x-2}{x}dx$

(c) $\int \frac{u^3+\sqrt[3]{2u}}{2u^2}du$

(d) $\int \frac{\left(\sqrt{t}-2\right)^2}{4t^2}dt$

Exercise 1.5.3 *Solve the initial value problems to find $f(t)$.*

(a) $f'(t)=3t+2$, *where* $f(0)=5$

(b) $f'(t)=\frac{t+3}{t}$, *where* $f(1)=-1$

(c) $f'(t)=\sin t$, *where* $f(\pi)=0$

(d) $f'(t)=e^{2x}-4$, *where* $f(0)=1$

Exercise 1.5.4 *A retailer determines that the marginal revenue function associated with selling q items is*

$$R'(q)=-0.008q+16$$

dollars per week per item.

(a) *Determine $R(q)$.*

(b) *Determine the demand equation relating the unit price p to the quantity q demanded.*

Exercise 1.5.5 *A supermarket determines that the marginal profit function associated with purchasing and selling q heads of lettuce is*

$$P'(q)=-0.003q+15,$$

dollars per year per unit. Given that the fixed storage cost is \$1000 per year, what is the maximum possible daily profit?

Exercise 1.5.6 *The same supermarket determines that the marginal cost of storing q bouquets of flowers is approximately*

$$C'(q)=0.001q+50$$

dollars per week, with a fixed cost of \$500 per week. Determine the total monthly cost of storing 500 bouquets a week.

Exercise 1.5.7 *A microbiologist is in charge of counting bacteria in a Petri dish. Starting with a single bacterium, she determines that the organisms are increasing at a rate of*

$$N'(t) = -2t^2 + 10t + 100$$

for $t \in [0,5]$ *hours.*

(a) *Determine as a function of* t *the number of organisms in the Petri dish.*

(b) *Over the entire interval studied, what was the average number of organisms in the dish?*

2. Techniques of Integration

Over the next few sections we examine some techniques that are frequently successful when seeking antiderivatives of functions.

2.1 Substitution Rule

2.1.1. Substitution Rule for Indefinite Integrals

Needless to say, most integration problems we will encounter will not be so simple. That is to say we will require more than the basic integration rules we have seen. Here's a slightly more complicated example: Find

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

This is not a "simple" derivative, but a little thought reveals that it must have come from an application of the Chain Rule. Multiplied on the "outside" is $2x$, which is the derivative of the "inside" function x^2. Checking:

$$\frac{d}{dx}\sin(x^2) = \cos(x^2)\frac{d}{dx}x^2 = 2x\cos(x^2)$$

so

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

To summarize: If we suspect that a given function is the derivative of another via the Chain Rule, we let u denote a likely candidate for the inner function, then translate the given function so that it is written entirely in terms of u, with no x remaining in the expression. If we can integrate this new function of u, then the antiderivative of the original function is obtained by replacing u by the equivalent expression in x. This result leads us to the following theorem.

Theorem 2.1: Substitution Rule for Indefinite Integrals

If $u = g(x)$ is a differentiable function whose range is an interval I and f is continuous on I, then

$$\int f(g(x))g'(x)\,dx = \int f(u)\,du.$$

We can describe two methods how the Substitution Rule may unfold in an integration process.

Method 1: Even in simple cases you may prefer to use this mechanical procedure, since it often helps to avoid silly mistakes. For example, consider again this simple problem:

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

Let $u = x^2$, then $du/dx = 2x$ or $du = 2x\,dx$. Since we have exactly $2x\,dx$ in the original integral, we can replace it by du:

$$\int 2x\cos(x^2)\,dx = \int \cos u\,du = \sin u + C = \sin(x^2) + C.$$

Method 2: This is not the only way to do the algebra, and typically there are many paths to the correct answer. For example, since $du/dx = 2x$, we have that $dx = du/2x$, and so the integral becomes

$$\int 2x\cos(x^2)\,dx = \int 2x\cos u\,\frac{du}{2x} = \int \cos u\,du.$$

Guideline for Substitution Rule

Given the integral

$$I = \int f(g(x))g'(x)\,dx$$

with f continuous and g differentiable, the following steps outline the **Substitution Rule process** for integrating I.

1. Let $u = g(x)$, which is typically the inside function of the function composition in the integrand.
2. Compute $du = g'(x)dx$.
3. Write the integral I solely in terms of u using either method:
   (a) use the substitution $u = g(x)$ and $du = g'(x)dx$; or
   (b) replace dx with $du/g'(x)$ and cancel $g'(x)$.
4. Integrate with respect to u.
5. Replace u by $g(x)$ to write the result in terms of x only.

Note:

1. However the Substitution Rule unfolds in the solution process, the important part is that you must eliminate all instances of the original variable, often x, during the process and then write your result back in terms of the original variable x.
2. Sometimes, the integrand has to be rearranged to see whether the Substitution Rule is a possible integration technique.
3. If a *first* substitution did not work out, then try to simplify or rearrange the integrand to see if a *different* substitution can be used.

4. As a general guideline for the Substitution Rule, we look for the *inside* function u to be

   - the radicand under a root: e.g.,

   $$\text{when } \int x^3\sqrt{x^2-5}\,dx \text{ we choose } u = x^2-5;$$

   - the base in a power with a real exponent: e.g.,

   $$\text{when } \int x\left(x^2-5\right)^5\,dx \text{ we choose } u = x^2-5;$$

   - the exponent in a power with a real base: e.g.,

   $$\text{when } \int x5^{x^2-5}\,dx \text{ we choose } u = x^2-5;$$

   - the denominator in a fraction: e.g.,

   $$\text{when } \int \frac{x}{x^2-5}\,dx \text{ we choose } u = x^2-5.$$

   However, sometimes u can be something different than is suggested above, so be open minded about this process.

Example 2.2: Substitution Rule

Evaluate $\int (ax+b)^n\,dx$, *assuming* a,b *are constants,* $a \neq 0$, *and* n *is a positive integer.*

Solution. We let $u = ax+b$ so $du = a\,dx$ or $dx = du/a$. Then

$$\int (ax+b)^n\,dx = \int \frac{1}{a}u^n\,du = \frac{1}{a(n+1)}u^{n+1} + C = \frac{1}{a(n+1)}(ax+b)^{n+1} + C.$$

♣

Example 2.3: Substitution Rule

Evaluate $\int \sin(ax+b)\,dx$, *assuming that* a *and* b *are constants and* $a \neq 0$.

Solution. Again we let $u = ax+b$ so $du = a\,dx$ or $dx = du/a$. Then

$$\int \sin(ax+b)\,dx = \int \frac{1}{a}\sin u\,du = \frac{1}{a}(-\cos u) + C = -\frac{1}{a}\cos(ax+b) + C.$$

♣

Example 2.4: Substitution in Denominator

Evaluate the following integral: $\displaystyle\int \frac{2y}{\sqrt{1-4y^2}}\,dy.$

Solution. We try the substitution:

$$u = 1-4y^2.$$

Then,

$$du = -8y\,dy$$

In the numerator we have $2y\,dy$, so rewriting the differential gives:

$$-\frac{1}{4}du = 2y\,dy.$$

Then the integral is:

$$\begin{aligned}
\int \frac{2y}{\sqrt{1-4y^2}}\,dy &= \int \left(1-4y^2\right)^{-1/2}(2y\,dy) \\
&= \int u^{-1/2}\left(-\frac{1}{4}du\right) \\
&= \left(\frac{-1}{4}\right)\frac{u^{1/2}}{1/2}+C \\
&= -\frac{\sqrt{1-4y^2}}{2}+C
\end{aligned}$$

♣

Example 2.5: Substitution in Base

Evaluate the following integral: $\displaystyle\int \cos x(\sin x)^5\,dx.$

Solution. In this question we will let $u = \sin x$. Then,

$$du = \cos x\,dx.$$

Thus, the integral becomes:

$$\begin{aligned}
\int \cos x(\sin x)^5\,dx &= \int u^5\,du \\
&= \frac{u^6}{6}+C \\
&= \frac{(\sin x)^6}{6}+C
\end{aligned}$$

♣

Example 2.6: Substitution

Evaluate the following integral: $\displaystyle\int \frac{\cos(\sqrt{x})}{\sqrt{x}}\,dx.$

Solution. We use the substitution:

$$u = x^{1/2}.$$

Then,

$$du = \frac{1}{2}x^{-1/2}\,dx.$$

Rewriting the differential we get:

$$2\,du = \frac{1}{\sqrt{x}}\,dx.$$

The integral becomes:

$$\begin{aligned}\int \frac{\cos(\sqrt{x})}{\sqrt{x}}\,dx &= 2\int \cos u\,du \\ &= 2\sin u + C \\ &= 2\sin(\sqrt{x}) + C\end{aligned}$$

♣

Example 2.7: Substitution

Evaluate the following integral: $\displaystyle\int 2x^3\sqrt{x^2+1}\,dx.$

Solution. This problem is a little bit different than the previous ones. It makes sense to let:

$$u = x^2 + 1,$$

then

$$du = 2x\,dx.$$

Making this substitution gives:

$$\begin{aligned}\int 2x^3\sqrt{x^2+1}\,dx &= \int x^2\sqrt{x^2+1}(2x)\,dx \\ &= \int x^2 u^{1/2}\,du\end{aligned}$$

This is a problem because our integrals can't have a mixture of two variables in them. Usually this means we chose our u incorrectly. However, in this case we can eliminate the remaining x's from our integral by using:

$$u = x^2 + 1 \implies x^2 = u - 1.$$

We get:

$$\int x^2 u^{1/2}\, du = \int (u-1)u^{1/2}\, du.$$

Now we proceed by simplifying the integrand and noticing we are left with power functions, which are readily integrated.

$$\int (u-1)u^{1/2}\, du = \int u^{3/2} - u^{1/2}\, du = \frac{2}{5}u^{5/2} - \frac{2}{3}u^{3/2} + C.$$

Therefore, our original integral becomes

$$\int 2x^3 \sqrt{x^2+1}\, dx = \frac{2}{5}(x^2+1)^{5/2} - \frac{2}{3}(x^2+1)^{3/2} + C$$

♣

Example 2.8: Two Choices for Substitution

Evaluate $\displaystyle\int \frac{2x}{\sqrt[3]{x^2-5}}\, dx.$

Solution.

Method 1: Let $u = x^2 - 5$, $du = 2x dx$.

$$\begin{aligned}
\int \frac{2x}{\sqrt[3]{x^2-5}}\, dx &= \int \frac{du}{u^{1/3}} \\
&= \frac{u^{2/3}}{2/3} + C \\
&= \frac{3}{2}u^{2/3} + C \\
&= \frac{3}{2}\left(x^2-5\right)^{2/3} + C
\end{aligned}$$

Method 2: Let $u = \sqrt[3]{x^2-5}$, $u^3 = x^2 - 5$, $3u^2 du = 2x dx$.

$$\begin{aligned}
\int \frac{2x}{\sqrt[3]{x^2-5}}\, dx &= \int \frac{3u^2 du}{u} \\
&= 3\int u du \\
&= 3\frac{u^2}{2} + C \\
&= \frac{3}{2}\left(x^2-5\right)^{2/3} + C
\end{aligned}$$

♣

Example 2.9: Application in Sales Projection

An airline wants to predict the number of tickets on a particular route that will be sold. They find that over the next year, ticket sales growth after t weeks can be modelled by

$$1000 - 500e^{-0.025t}$$

tickets per week for $t \in [0, 52]$.

(a) Determine the predicted total number of tickets sold after the first t weeks.

(b) Determine the predicted total number of tickets sold in the next year.

Solution.

(a) Let $N(t)$ be the total number of tickets predicted to be sold after the first t months. We are given

$$N'(t) = 1000 - 500e^{-0.025t}.$$

Thus,

$$N(t) = \int \left(1000 - 500e^{-0.025t}\right) dt.$$

We integrate by substitution. Let $u = -0.025t$, $du = -0.025dt$:

$$\begin{aligned} N(t) &= \int \left(1000 - 500e^{-0.025t}\right) dt \\ &= \int 1000dt + \frac{500}{0.025} \int e^u\, du \\ &= 1000t + 20{,}000e^{-0.025t} + C \end{aligned}$$

It remains to determine the value of C. We require

$$N(0) = 0 \implies C = -20{,}000.$$

Therefore, the total number of tickets sold after the first t months is predicted to be

$$N(t) = 1000t + 20{,}000e^{-0.025t} - 20{,}000.$$

(b) Since

$$N(52) = 1000(52) + 20{,}000e^{-0.025(52)} - 20{,}000 = 37450.64,$$

we have that the model predicts that 37,451 tickets will be sold on this route in the next year.

♣

2.1.2. Substitution Rule for Definite Integrals

The next example shows how to use the Substitution Rule when dealing with definite integrals.

Example 2.10: Substitution Rule with Two Methods

Evaluate $\int_2^4 x\sin(x^2)\,dx$.

Solution.
Method 1: First Compute the Antiderivative, then Evaluate the Definite Integral.
Let $u = x^2$, so $du = 2x\,dx$ or $x\,dx = du/2$. Then

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

Now

$$\int_2^4 x\sin(x^2)\,dx = -\frac{1}{2}\cos(x^2)\Big|_2^4 = -\frac{1}{2}\cos(16) + \frac{1}{2}\cos(4).$$

Method 2: Changing the Limits of Integration with the Substitution.
A somewhat neater alternative to this method is to change the original limits to match the variable u. Since $u = x^2$, when $x = 2$, $u = 4$, and when $x = 4$, $u = 16$. So we can do this:

$$\int_2^4 x\sin(x^2)\,dx = \int_4^{16} \frac{1}{2}\sin u\,du = -\frac{1}{2}(\cos u)\Big|_4^{16} = -\frac{1}{2}\cos(16) + \frac{1}{2}\cos(4).$$

An incorrect, and dangerous, alternative is something like this:

$$\int_2^4 x\sin(x^2)\,dx = \int_2^4 \frac{1}{2}\sin u\,du = -\frac{1}{2}\cos(u)\Big|_2^4 = -\frac{1}{2}\cos(x^2)\Big|_2^4 = -\frac{1}{2}\cos(16) + \frac{1}{2}\cos(4).$$

This is incorrect because $\int_2^4 \frac{1}{2}\sin u\,du$ means that u takes on values between 2 and 4, which is wrong. It is dangerous, because it is very easy to get to the point $-\frac{1}{2}\cos(u)\Big|_2^4$ and forget to substitute x^2 back in for u, thus getting the incorrect answer $-\frac{1}{2}\cos(4) + \frac{1}{2}\cos(2)$. An acceptable alternative is to clearly indicate that the limit substitution is to be done with respect to the x variable, using something like:

$$\int_2^4 x\sin(x^2)\,dx = \int_{x=2}^{x=4} \frac{1}{2}\sin u\,du = -\frac{1}{2}\cos(u)\Big|_{x=2}^{x=4} = -\frac{1}{2}\cos(x^2)\Big|_2^4 = -\frac{\cos(16)}{2} + \frac{\cos(4)}{2}.$$

♣

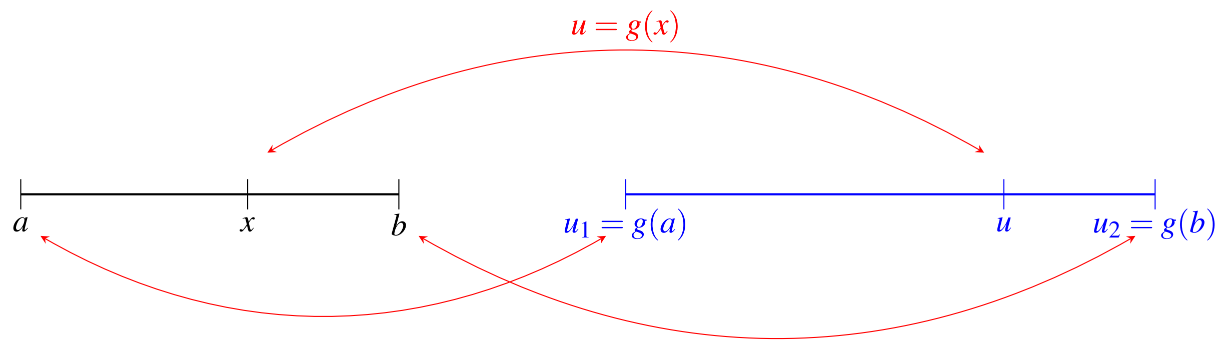
To summarize, we have the following.

Theorem 2.11: Substitution Rule for Definite Integrals

If g' is continuous on $[a,b]$ and f is continuous on the range of $u = g(x)$, then

$$\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)\,du.$$

Note: When using the Substitution Rule for integrating definite integrals, it is important to change the limits of integration from those of the original function to those of the substituted function. Otherwise, the definite integral will evaluate to an incorrect result.



Example 2.12: Substitution Rule for Definite Integrals

Evaluate $\displaystyle\int_{1/4}^{1/2} \frac{\cos(\pi t)}{\sin^2(\pi t)}\,dt.$

Solution. Let $u = \sin(\pi t)$ so $du = \pi\cos(\pi t)\,dt$ or $du/\pi = \cos(\pi t)\,dt$. We change the limits to $u(1/4) = \sin(\pi/4) = \sqrt{2}/2$ and $u(1/2) = \sin(\pi/2) = 1$. Then

$$\int_{1/4}^{1/2} \frac{\cos(\pi t)}{\sin^2(\pi t)}\,dt = \int_{\sqrt{2}/2}^{1} \frac{1}{\pi}\frac{1}{u^2}\,du = \int_{\sqrt{2}/2}^{1} \frac{1}{\pi}u^{-2}\,du = \frac{1}{\pi}\frac{u^{-1}}{-1}\bigg|_{\sqrt{2}/2}^{1} = -\frac{1}{\pi} + \frac{\sqrt{2}}{\pi}.$$

♣

Example 2.13: Substitution Rule for Definite Integrals

Evaluate $\displaystyle\int_{-1}^{1} (y+1)\left(y^2+2y\right)^8 dy.$

Solution. Although we could expand the integrand, since this would yield powers of y which we can certainly integrate without using the Substitution Rule at all, the exponent 8 would make this a rather messy process that is surely prone to errors. Instead we proceed with the obvious choice of substitution and let $u = y^2 + 2y$, then

$$dy = (2y+2)dy = 2(y+1)dy \implies \frac{du}{2} = (y+1)dy$$

with limits of integration

$$u(-1) = (-1)^2 + 2(-1) = -1 \text{ and } u(1) = (1)^2 + 2(1) = 3.$$

Take note that sometimes the value of a limit of integration does not change. The point is that one still needs to substitute the values of the original variable of integration, in this case y, to work with the limits of integration for the substituted variable, in this case u.

Then the integral evaluates as follows

$$\int_{-1}^{1} (y+1)\left(y^2+2y\right)^8 dy = \int_{-1}^{3} \frac{u^8}{2} du = \left.\frac{u^9}{18}\right|_{-1}^{3} = \frac{3^9}{18} - \frac{(-1)^9}{18} = \frac{19,684}{18}.$$

♣

Exercises for Section 2.1

Exercise 2.1.1 *Evaluate the following indefinite integrals.*

(a) $\displaystyle\int (1-t)^9\, dt$

(b) $\displaystyle\int (x^2+1)^2\, dx$

(c) $\displaystyle\int x(x^2+1)^{100}\, dx$

(d) $\displaystyle\int \frac{1}{\sqrt[3]{1-5t}}\, dt$

(e) $\displaystyle\int \sin^3 x \cos x\, dx$

(f) $\displaystyle\int x\sqrt{100-x^2}\, dx$

(g) $\displaystyle\int \frac{x^2}{\sqrt{1-x^3}}\, dx$

(h) $\displaystyle\int \cos(\pi t)\cos\big(\sin(\pi t)\big)\, dt$

(i) $\displaystyle\int \frac{\sin x}{\cos^3 x}\, dx$

(j) $\displaystyle\int \tan x\, dx$

(k) $\displaystyle\int \sec^2 x \tan x\, dx$

(l) $\displaystyle\int \frac{\sin(\tan x)}{\cos^2 x}\, dx$

(m) $\displaystyle\int \frac{6x}{(x^2-7)^{1/9}}\, dx$

(n) $\displaystyle\int f(x)f'(x)\, dx$

Exercise 2.1.2 *Evaluate the following definite integrals.*

(a) $\displaystyle\int_0^{\pi} \sin^5(3x)\cos(3x)\, dx$

(b) $\displaystyle\int_0^{\sqrt{\pi}/2} x\sec^2(x^2)\tan(x^2)\, dx$

(c) $\displaystyle\int_3^4 \frac{1}{(3x-7)^2}\, dx$

(d) $\displaystyle\int_0^{\pi/6} (\cos^2 x - \sin^2 x)\, dx$

(e) $\displaystyle\int_{-1}^{1}(2x^3-1)(x^4-2x)^6\,dx$

(f) $\displaystyle\int_{-1}^{1}\sin^7 x\,dx$

Exercise 2.1.3 *A toy truck manufacturer estimates that the number of sales after Christmas declines at a rate of*

$$-3e^{-0.2t}$$

toys per day ($t \in [0,60]$). If the manufacturer sells 10,000 units on Christmas day ($t=0$), determine the number of expected sales after t days.

Exercise 2.1.4 *Let q be the quantity (in thousands) of a product in the market when the unit price is set at p dollars per unit. There are currently 2000 units available at a price of \$2/unit. Determine the corresponding supply equation if the price increases at a rate of*

$$p'(q) = \frac{100q}{(1-q)^2}.$$

Exercise 2.1.5 *Let q represent the daily quantity demanded (in thousands) of a certain product. When the unit price p is set at \$3, it is observed that 1000 units are demanded per day. Determine the corresponding demand equation if the price decreases at a rate of*

$$p'(q) = \frac{-200q}{(3+q^2)^{3/2}}.$$

Exercise 2.1.6 *In the first month ($t=0$) following a successful marketing campaign, a company sells 2,000 units of their product. Management predicts that the sales will decline at a rate of*

$$N'(t) = -200(e^{-2t}+1)$$

units per month for the next 3 months. Determine the total number of expected sales after 3 months.

2.2 Powers of Trigonometric Functions

The trigonometric substitutions we will focus on in this section are summarized in the table below:

Substitution	$u=\sin x$	$u=\cos x$	$u=\tan x$	$u=\sec x$
Derivative	$du=\cos x\,dx$	$du=-\sin x\,dx$	$du=\sec^2 x\,dx$	$du=\sec x\tan x\,dx$

2.2.1. Products of Powers of Sine and Cosine

Functions consisting of powers of the sine and cosine can be integrated by using substitution and trigonometric identities. These can sometimes be tedious, but the technique is straightforward. A similar technique is applicable to powers of secant and tangent as shown in Section 2.2.2 (and also cosecant and cotangent, not discussed here). An example will suffice to explain the approach.

Example 2.14: Odd Power of Sine

Evaluate $\displaystyle\int \sin^5 x\,dx$.

Solution. Rewrite the function:

$$\begin{aligned}
\int \sin^5 x\,dx &= \int \sin x \sin^4 x\,dx \\
&= \int \sin x(\sin^2 x)^2\,dx \\
&= \int \sin x(1-\cos^2 x)^2\,dx.
\end{aligned}$$

Now use $u = \cos x$, $du = -\sin x\,dx$:

$$\begin{aligned}
\int \sin x(1-\cos^2 x)^2\,dx &= \int -(1-u^2)^2\,du \\
&= \int -(1-2u^2+u^4)\,du \\
&= \int 1+2u^2-u^4\,du \\
&= -u+\frac{2}{3}u^3-\frac{1}{5}u^5+C \\
&= -\cos x+\frac{2}{3}\cos^3 x-\frac{1}{5}\cos^5 x+C.
\end{aligned}$$

♣

Observe that by taking the substitution $u = \cos x$ in the last example, we ended up with an even power of sine from which we can use the formula $\sin^2 x + \cos^2 x = 1$ to replace any remaining sines. We then ended up with a polynomial in u in which we could expand and integrate quite easily. Notice here that the "obvious" substitution $u = \sin x$ in the original integral does not lead to any useful simplification.

This technique works for products of powers of sine and cosine. We summarize it below.

Guideline for Integrating Products of Sine and Cosine

When evaluating $\int \sin^m x \cos^n x \, dx$:

1. **The power of sine is odd (m odd):**
   (a) Use $u = \cos x$ and $du = -\sin x \, dx$.
   (b) Replace dx using (a), thus cancelling one power of $\sin x$ by the substitution of du, and be left with an even number of sine powers.
   (c) Use $\sin^2 x = 1 - \cos^2 x$ $(= 1 - u^2)$ to replace the leftover sines.
2. **The power of cosine is odd (n odd):**
   (a) Use $u = \sin x$ and $du = \cos x \, dx$.
   (b) Replace dx using (a), thus cancelling one power of $\cos x$ by the substitution of du, and be left with an even number of cosine powers.
   (c) Use $\cos^2 x = 1 - \sin^2 x$ $(= 1 - u^2)$ to replace the leftover cosines.
3. **Both m and n are odd:**
   Use either 1 or 2 (both will work).
4. **Both m and n are even:**
   Use $\cos^2 x = \frac{1}{2}(1 + \cos(2x))$ and/or $\sin^2 x = \frac{1}{2}(1 - \cos(2x))$ to reduce to a form that can be integrated.

Note: As m and n get large, multiple steps will be needed.

Example 2.15: Odd Power of Cosine and Even Power of Sine

Evaluate $\int \sin^6 x \cos^5 x \, dx$.

Solution. We will show this solution in two ways.
Method 1: We rewrite the integrand to use the Substitution Rule.

$$\begin{aligned} \int \sin^6 x \cos^5 x \, dx &= \int \sin^6 x \cos^4 x \cos x \, dx \\ &= \int \sin^6 x (\cos^2 x)^2 \cos x \, dx \\ &= \int \sin^6 x (1 - \sin^2 x)^2 \cos x \, dx \end{aligned}$$

Then use the substitution $u = \sin x$ since the derivative of $\sin x$ is $\cos x$, and also because there is only one occurrence of cosine in the integrand. Hence, $du = \cos x \, dx$ is the perfect substitution.

$$\int \sin^6 x (1 - \sin^2 x)^2 \cos x \, dx = \int u^6 (1 - u^2)^2 \, du$$

$$\begin{aligned}
&= \int u^6 - 2u^8 + u^{10}\, du \\
&= \frac{u^7}{7} - \frac{2u^9}{9} + \frac{u^{11}}{11} + C \\
&= \frac{\sin^7 x}{7} - \frac{2\sin^9 x}{9} + \frac{\sin^{11} x}{11} + C
\end{aligned}$$

Method 2: We apply the *Guideline for Integrating Products of Sine and Cosine*. Since the power of cosine is odd, we use the substitution $u = \sin x$ and $du = \cos x$.

Then

$$\begin{aligned}
\int \sin^6 x \cos^5 x\, dx &= \int u^6 \cos^5 x \frac{du}{\cos x} && \text{Using the substitution} \\
&= \int u^6 \left(\cos^2 x\right)^2 du && \text{Canceling a } \cos x \text{ and rewriting } \cos^4 x \\
&= \int u^6 (1 - \sin^2 x)^2\, du && \text{Using trig identity } \cos^2 x = 1 - \sin^2 x \\
&= \int u^6 (1 - u^2)^2\, du && \text{Writing integral in terms of } u\text{'s} \\
&= \int u^6 - 2u^8 + u^{10}\, du && \text{Expand and collect like terms} \\
&= \frac{u^7}{7} - \frac{2u^9}{9} + \frac{u^{11}}{11} + C && \text{Integrating} \\
&= \frac{\sin^7 x}{7} - \frac{2\sin^9 x}{9} + \frac{\sin^{11} x}{11} + C && \text{Replacing } u \text{ back in terms of } x
\end{aligned}$$

♣

Example 2.16: Odd Power of Cosine

Evaluate $\int \cos^3 x\, dx$.

Solution. Since the power of cosine is odd, we use the substitution $u = \sin x$ and $du = \cos x\, dx$. This may seem strange at first since we don't have $\sin x$ in the question, but it does work!

$$\begin{aligned}
\int \cos^3 x\, dx &= \int \cos^3 x \frac{du}{\cos x} && \text{Using the substitution} \\
&= \int \cos^2 x\, du && \text{Canceling a } \cos x \\
&= \int (1 - \sin^2 x)\, du && \text{Using trig identity } \cos^2 x = 1 - \sin^2 x \\
&= \int (1 - u^2)\, du && \text{Writing integral in terms of } u\text{'s} \\
&= u - \frac{u^3}{3} + C && \text{Integrating} \\
&= \sin x - \frac{\sin^3 x}{3} + C && \text{Replacing } u \text{ back in terms of } x
\end{aligned}$$

♣

Example 2.17: Product of Even Powers of Sine and Cosine

Evaluate $\int \sin^2 x \cos^2 x\, dx$.

Solution. Use the formulas

$$\sin^2 x = \frac{1-\cos(2x)}{2} \quad \text{and} \quad \cos^2 x = \frac{1+\cos(2x)}{2}$$

to get

$$\int \sin^2 x \cos^2 x\, dx = \int \frac{1-\cos(2x)}{2} \cdot \frac{1+\cos(2x)}{2}\, dx.$$

We then have

$$\begin{aligned} \int \sin^2 x \cos^2 x\, dx &= \int \frac{1-\cos(2x)}{2} \cdot \frac{1+\cos(2x)}{2}\, dx \\ &= \frac{1}{4} \int 1 - \cos^2 2x\, dx \\ &= \frac{1}{4} \left(x - \int \cos^2 2x\, dx \right). \end{aligned}$$

To continue the integration, we use the cosine double angle identity again with

$$\cos^2(2x) = \frac{1+\cos(4x)}{2}.$$

Then

$$\begin{aligned} \int \sin^2 x \cos^2 x\, dx &= \frac{1}{4} \left(x - \frac{1}{2} \int 1 + \cos 4x\, dx \right) \\ &= \frac{1}{4} \left(x - \frac{1}{2} \left(x + \frac{\sin 4x}{4} \right) \right) \\ &= \frac{1}{4} \left(x - \frac{x}{2} - \frac{\sin 4x}{8} \right) + C \end{aligned}$$

♣

Example 2.18: Even Power of Sine

Evaluate $\int \sin^6 x\, dx$.

Solution. Use $\sin^2 x = (1-\cos(2x))/2$ to rewrite the function:

$$\int \sin^6 x\, dx \quad = \quad \int (\sin^2 x)^3\, dx$$

$$\begin{aligned} &= \int \frac{(1-\cos 2x)^3}{8}\,dx \\ &= \frac{1}{8}\int 1 - 3\cos 2x + 3\cos^2 2x - \cos^3 2x\,dx. \end{aligned}$$

Now we have four integrals to evaluate. Ignoring the constant for now:

$$\int 1\,dx = x$$

and

$$\int -3\cos 2x\,dx = -\frac{3}{2}\sin 2x$$

are easy. The $\cos^3 2x$ integral is like the previous example:

$$\begin{aligned} \int -\cos^3 2x\,dx &= \int -\cos 2x \cos^2 2x\,dx \\ &= \int -\cos 2x(1-\sin^2 2x)\,dx \\ &= \int -\frac{1}{2}(1-u^2)\,du \\ &= -\frac{1}{2}\left(u - \frac{u^3}{3}\right) \\ &= -\frac{1}{2}\left(\sin 2x - \frac{\sin^3 2x}{3}\right). \end{aligned}$$

And finally we use another trigonometric identity, $\cos^2 x = (1+\cos(2x))/2$:

$$\int 3\cos^2 2x\,dx = 3\int \frac{1+\cos 4x}{2}\,dx = \frac{3}{2}\left(x + \frac{\sin 4x}{4}\right).$$

So at long last we get

$$\int \sin^6 x\,dx = \frac{x}{8} - \frac{3}{16}\sin 2x - \frac{1}{16}\left(\sin 2x - \frac{\sin^3 2x}{3}\right) + \frac{3}{16}\left(x + \frac{\sin 4x}{4}\right) + C.$$

♣

2.2.2. Exploring Powers of Secant and Tangent

Now, we turn our attention to products of secant and tangent. Some we already know how to do from the table of Integral Rules in Section 1.5.3.

$$\int \sec^2 x\,dx = \tan x + C \qquad\qquad \int \sec x \tan x\,dx = \sec x + C$$

Note: A common mistake is to believe that $\int \tan x\,dx$ is $\sec^2(x) + C$ – this is *not* true. We can readily integrate $\tan x$ to demonstrate why this is not true.

Example 2.19: Integrating Tangent

Evaluate $\displaystyle\int \tan x\, dx$.

Solution. Note that $\tan x = \dfrac{\sin x}{\cos x}$ and let $u = \cos x$, so that $du = -\sin x\, dx$.

$$\begin{aligned}
\int \tan x\, dx &= \int \frac{\sin x}{\cos x}\, dx && \text{Rewriting } \tan x\\
&= \int \frac{\sin x}{u} \frac{du}{-\sin x} && \text{Using the substitution}\\
&= -\int \frac{1}{u}\, du && \text{Cancelling and pulling the } -1 \text{ out}\\
&= -\ln|u| + C && \text{Using formula } \int \frac{1}{u}\, dx = \ln|u| + C\\
&= -\ln|\cos x| + C && \text{Replacing } u \text{ back in terms of } x\\
&= \ln|\sec x| + C && \text{Using log properties and } \sec x = 1/\cos x
\end{aligned}$$

♣

Example 2.20: Integrating Tangent Squared

Evaluate $\displaystyle\int \tan^2 x\, dx$.

Solution. Note that $\tan^2 x = \sec^2 x - 1$.

$$\begin{aligned}
\int \tan^2 x\, dx &= \int \left(\sec^2 x - 1\right)\, dx && \text{Rewriting } \tan x\\
&= \tan x - x + C && \text{Since } \int \sec^2 x\, dx = \tan x + C
\end{aligned}$$

♣

Note: In problems with tangent and secant, two integrals come up frequently:

$$\int \sec^3 x\, dx \qquad \text{and} \qquad \int \sec x\, dx.$$

Both have relatively nice expressions but they are a bit tricky to discover.

First we evaluate $\int \sec x\, dx$, which we will need to compute $\int \sec^3 x\, dx$.

Example 2.21: Integral of Secant

Evaluate $\displaystyle\int \sec x\, dx$.

Solution. We start with a trick, namely we multiply the integrand by 1, but we express 1 as the ratio $(\sec x+\tan x)/(\sec x+\tan x)$. This sounds like a crazy idea, but it works!

$$\begin{aligned}\int \sec x\,dx &= \int \sec x\frac{\sec x+\tan x}{\sec x+\tan x}\,dx\\ &= \int \frac{\sec^2 x+\sec x\tan x}{\sec x+\tan x}\,dx.\end{aligned}$$

Now let $u=\sec x+\tan x$, $du=\sec x\tan x+\sec^2 x\,dx$, exactly the numerator of the function we are integrating. Thus

$$\begin{aligned}\int \sec x\,dx &= \int \frac{\sec^2 x+\sec x\tan x}{\sec x+\tan x}\,dx\\ &= \int \frac{1}{u}\,du=\ln|u|+C\\ &= \ln|\sec x+\tan x|+C.\end{aligned}$$

♣

Now we compute the integral $\int \sec^3 x\,dx$.

Example 2.22: Integral of Secant Cubed

Evaluate $\int \sec^3 x\,dx$.

Solution.

$$\begin{aligned}\sec^3 x &= \frac{\sec^3 x}{2}+\frac{\sec^3 x}{2}=\frac{\sec^3 x}{2}+\frac{(\tan^2 x+1)\sec x}{2}\\ &= \frac{\sec^3 x}{2}+\frac{\sec x\tan^2 x}{2}+\frac{\sec x}{2}\\ &= \frac{\sec^3 x+\sec x\tan^2 x}{2}+\frac{\sec x}{2}.\end{aligned}$$

We already know how to integrate $\sec x$, so we just need the first quotient. This is "simply" a matter of recognizing the Product Rule differentiation in action:

$$\int \sec^3 x+\sec x\tan^2 x\,dx=\sec x\tan x.$$

So putting these together we get

$$\int \sec^3 x\,dx=\frac{\sec x\tan x}{2}+\frac{\ln|\sec x+\tan x|}{2}+C,$$

♣

2.2.3. Products of Powers of Secant and Tangent

For products of secant and tangent it is best to use the following Guideline.

Guideline for Integrating Products of Secant and Tangent

When evaluating $\int \sec^m x \tan^n x\, dx$:

1. **The power of secant is even (m even):**
   (a) Use $u = \tan x$ and $du = \sec^2 x\, dx$.
   (b) Cancel $\sec^2 x$ by the substitution of dx, and be left with an even number of secants.
   (c) Use $\sec^2 x = 1 + \tan^2 x$ $(= 1 + u^2)$ to replace the leftover secants.
2. **The power of tangent is odd (n odd):**
   (a) Use $u = \sec x$ and $du = \sec x \tan x\, dx$.
   (b) Cancel one $\sec x$ and one $\tan x$ by the substitution of dx. The number of remaining tangents is even.
   (c) Use $\tan^2 x = \sec^2 x - 1$ $(= u^2 - 1)$ to replace the leftover tangents.
3. **m is even or n is odd:**
   Use either 1 or 2 (both will work).
4. **The power of secant is odd and the power of tangent is even:**
   No guideline. The integrals $\int \sec x\, dx$ and $\int \sec^3 x\, dx$ can usually be looked up, or recalled from memory.

Example 2.23: Even Power of Secant

Evaluate $\int \sec^6 x \tan^6 x\, dx$.

Solution. Since the power of secant is even, we use $u = \tan x$, so that $du = \sec^2 x\,dx$.

$$\begin{aligned}
\int \sec^6 x \tan^6 x\,dx &= \int \sec^6 x\,(u^6)\,\frac{du}{\sec^2 x} && \text{Using the substitution}\\
&= \int \sec^4 x(u^6)\,du && \text{Cancelling a } \sec^2 x\\
&= \int (\sec^2 x)^2(u^6)\,du && \text{Rewriting } \sec^4 x\\
&= \int (1+\tan^2 x)^2(u^6)\,du && \text{Using } \sec^2 x = 1+\tan^2 x\\
&= \int (1+u^2)^2(u^6)\,du && \text{Using the substitution}
\end{aligned}$$

To integrate this product the easiest method is expand it into a polynomial and integrate term-by-term.

$$\begin{aligned}
\int \sec^6 x \tan^6 x\,dx &= \int (u^6+2u^8+u^{10})\,du && \text{Expanding}\\
&= \frac{u^7}{7}+\frac{2u^9}{9}+\frac{u^{11}}{11}+C && \text{Integrating}\\
&= \frac{\tan^7 x}{7}+\frac{2\tan^9 x}{9}+\frac{\tan^{11} x}{11}+C && \text{Rewriting in terms of } x
\end{aligned}$$

♣

Example 2.24: Odd Power of Tangent

Evaluate $\int \sec^5 x \tan x\,dx$.

Solution. Since the power of tangent is odd, we use $u = \sec x$, so that $du = \sec x \tan x\,dx$. Then we have:

$$\begin{aligned}
\int \sec^5 x \tan x\,dx &= \int \sec^5 x \tan x\,\frac{du}{\sec x \tan x} && \text{Substituting } dx \text{ first}\\
&= \int \sec^4 x\,du && \text{Cancelling}\\
&= \int u^4\,du && \text{Using the substitution}\\
&= \frac{u^5}{5}+C && \text{Integrating}\\
&= \frac{\sec^5 x}{5}+C && \text{Rewriting in terms of } x
\end{aligned}$$

♣

Example 2.25: Odd Power of Secant and Even Power of Tangent

Evaluate $\int \sec x \tan^2 x\,dx$.

Solution. The Guideline doesn't help us in this scenario. However, since $\tan^2 x = \sec^2 x - 1$, we have

$$\begin{aligned}
\int \sec x \tan^2 x\,dx &= \int \sec x(\sec^2 x - 1)\,dx \\
&= \int (\sec^3 x - \sec x)\,dx \\
&= \frac{1}{2}(\sec x \tan x + \ln|\sec x + \tan x|) - \ln|\sec x + \tan x| + C \\
&= \frac{1}{2}\sec x \tan x + \frac{1}{2}\ln|\sec x + \tan x| - \ln|\sec x + \tan x| + C \\
&= \frac{1}{2}\sec x \tan x - \frac{1}{2}\ln|\sec x + \tan x| + C
\end{aligned}$$

♣

Exercises for Section 2.2

Exercise 2.2.1 *Evaluate the following indefinite integrals.*

(a) $\int \sin^2 x\,dx$

(b) $\int \sin^3 x\,dx$

(c) $\int \sin^4 x\,dx$

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

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

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

(g) $\int \sin x(\cos x)^{3/2}\,dx$

(h) $\int \sec^2 x \csc^2 x\,dx$

(i) $\int \tan^3 x \sec x\,dx$

(j) $\int \left(\frac{1}{\csc x} + \frac{1}{\sec x}\right) dx$

(k) $\int \frac{\cos^2 x + \cos x + 1}{\cos^3 x}\,dx$

(l) $\int x\sec^2(x^2)\tan^4(x^2)\,dx$

(m) $\int x\sec^2(x^2)\tan^4(x^2)\,dx$

2.3 Trigonometric Substitutions

So far we have seen that it sometimes helps to replace a subexpression of a function by a single variable. Occasionally it can help to replace the original variable by something more complicated. This seems like a "reverse" substitution, but it is really no different in principle than ordinary substitution.

Example 2.26: Sine Substitution

Evaluate $\int \sqrt{1-x^2}\,dx.$

Solution. Let $x = \sin u$ so $dx = \cos u\,du$. Then

$$\int \sqrt{1-x^2}\,dx = \int \sqrt{1-\sin^2 u}\cos u\,du = \int \sqrt{\cos^2 u}\cos u\,du.$$

We would like to replace $\sqrt{\cos^2 u}$ by $\cos u$, but this is valid only if $\cos u$ is positive, since $\sqrt{\cos^2 u}$ is positive. Consider again the substitution $x = \sin u$. We could just as well think of this as $u = \arcsin x$. If we do, then by the definition of the arcsine, $-\pi/2 \le u \le \pi/2$, so $\cos u \ge 0$ and so we are allowed to continue and perform the simplification:

$$\begin{aligned}\int \sqrt{\cos^2 u}\cos u\,du &= \int \cos^2 u\,du \\ &= \int \frac{1+\cos 2u}{2}\,du = \frac{u}{2} + \frac{\sin 2u}{4} + C \\ &= \frac{\arcsin x}{2} + \frac{\sin(2\arcsin x)}{4} + C.\end{aligned}$$

This is a perfectly good answer, though the term $\sin(2\arcsin x)$ is a bit unpleasant. It is possible to simplify this. Using the identity $\sin 2x = 2\sin x\cos x$, we can write $\sin 2u = 2\sin u\cos u = 2\sin(\arcsin x)\sqrt{1-\sin^2 u} = 2x\sqrt{1-\sin^2(\arcsin x)} = 2x\sqrt{1-x^2}$. Then the full antiderivative is

$$\frac{\arcsin x}{2} + \frac{2x\sqrt{1-x^2}}{4} + C = \frac{\arcsin x}{2} + \frac{x\sqrt{1-x^2}}{2} + C.$$

♣

Note:

1. This type of substitution is usually indicated when the function you wish to integrate contains a polynomial expression that might allow you to use the fundamental identity $\sin^2 x + \cos^2 x = 1$ in one of three forms:

   $$\cos^2 x = 1 - \sin^2 x \qquad \sec^2 x = 1 + \tan^2 x \qquad \tan^2 x = \sec^2 x - 1.$$

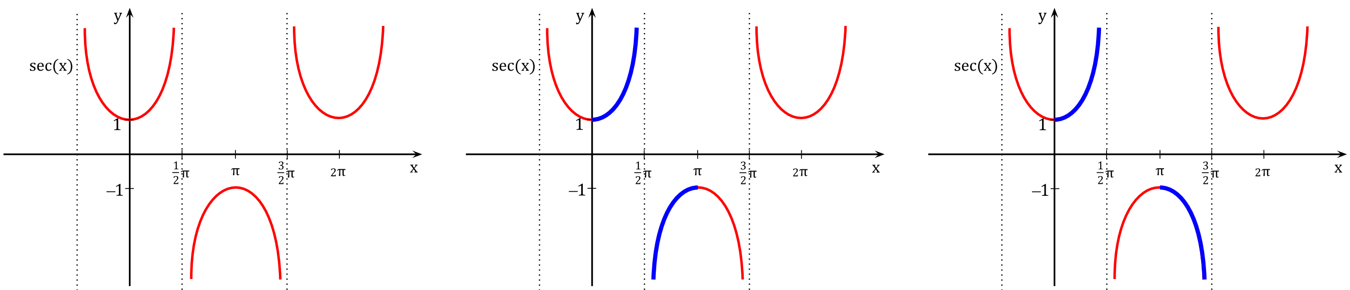
   If your function contains $1 - x^2$, as in the example above, try $x = \sin u$; if it contains $1 + x^2$ try $x = \tan u$; and if it contains $x^2 - 1$, try $x = \sec u$. Sometimes you will need to try something a bit different to handle constants other than *inverse substitution*, which is described next.

2. In a **traditional substitution** we let $u = u(x)$, i.e., our new variable is defined in terms of x. In an **inverse substitution** we let $x = g(u)$, i.e., we assume x can be written in terms of u. We cannot do this arbitrarily since we do **NOT** get to "choose" x. For example, an inverse substitution of $x = 1$ will give an obviously wrong answer. However, when $x = g(u)$ is an invertible function, then we are really doing a u-substitution with $u = g^{-1}(x)$. Now the Substitution Rule applies.

3. Sometimes with inverse substitutions involving trig functions we use θ instead of u. Thus, we would take $x = \sin\theta$ instead of $x = \sin u$.

4. We would like our inverse substitution $x = g(u)$ to be a one-to-one function, and $x = \sin u$ is not one-to-one. In the next few paragraphs, we discuss how we can overcome this issue by using the restricted trigonometric functions.

The three common **trigonometric substitutions** are the restricted sine, restricted tangent and restricted secant. Thus, for sine we use the domain $[-\pi/2,\ \pi/2]$ and for tangent we use $(-\pi/2,\ \pi/2)$. Depending on the convention chosen, the restricted secant function is usually defined in one of two ways.



One convention is to restrict secant to the region $[0,\ \pi/2)\cup(\pi/2,\pi]$ as shown in the middle graph. The other convention is to use $[0,\ \pi/2)\cup[\pi,\ 3\pi/2)$ as shown in the right graph. Both choices give a one-to-one restricted secant function and no universal convention has been adopted. To make the analysis in this section less cumbersome, we will use the domain $[0,\ \pi/2)\cup[\pi,\ 3\pi/2)$ for the restricted secant function. Then $\sec^{-1}x$ is defined to be the inverse of this restricted secant function.

Typically trigonometric substitutions are used for problems that involve radical expressions. The table below outlines when each substitution is typically used along with their restricted intervals.

Expression	**Substitution**	**Restricted Interval**
$\sqrt{a^2-x^2}$	$x=a\sin\theta$	$\theta\in[-\pi/2,\ \pi/2]$
$\sqrt{a^2+x^2}$ or a^2+x^2	$x=a\tan\theta$	$\theta\in(-\pi/2,\ \pi/2)$
$\sqrt{x^2-a^2}$	$x=a\sec\theta$	$\theta\in[0,\ \pi/2)\cup[\pi,\ 3\pi/2)$

All three substitutions are one-to-one on the listed intervals. When dealing with radicals we often end up with absolute values since

$$\sqrt{z^2}=|z|.$$

For each of the three trigonometric substitutions above we will verify that we can ignore the absolute value in each case when encountering a radical.

For $x=a\sin\theta$, the expression $\sqrt{a^2-x^2}$ becomes

$$\sqrt{a^2-x^2}=\sqrt{a^2-a^2\sin^2\theta}=\sqrt{a^2(1-\sin^2\theta)}=a\sqrt{\cos^2\theta}=a|\cos\theta|=a\cos\theta$$

This is because $\cos\theta\geq 0$ when $\theta\in[-\pi/2,\ \pi/2]$. For $x=a\tan\theta$, the expression $\sqrt{a^2+x^2}$ becomes

$$\sqrt{a^2+x^2}=\sqrt{a^2+a^2\tan^2\theta}=\sqrt{a^2(1+\tan^2\theta)}=a\sqrt{\sec^2\theta}=a|\sec\theta|=a\sec\theta$$

This is because $\sec\theta>0$ when $\theta\in(-\pi/2,\ \pi/2)$.

Finally, for $x = a\sec\theta$, the expression $\sqrt{x^2-a^2}$ becomes

$$\sqrt{x^2-a^2} = \sqrt{a^2\sec^2\theta - a^2} = \sqrt{a^2(\sec^2\theta - 1)} = a\sqrt{\tan^2\theta} = a|\tan\theta| = a\tan\theta$$

This is because $\tan\theta \geq 0$ when $\theta \in [0,\ \pi/2)\cup[\pi,\ 3\pi/2)$.

Thus, when using an appropriate trigonometric substitution we can usually ignore the absolute value. After integrating, we typically get an answer in terms of θ (or u) and need to convert back to x's. To do so, we use the guideline below:

- For trig functions containing θ, use a triangle to convert to x's.
- For θ by itself, use the inverse trig function.

All pieces needed for such a trigonometric substitution can be summarized as follows:

Guideline for Trigonometric Substitution

Suppose we have an integral with any of the following expressions, then use the substitution, differential, identity and inverse of substitution listed below to guide yourself through the integration process:

Expression	**Substitution**	**Differential**	**Identity**	**Inverse of Substitution**
$\sqrt{a^2-x^2}$	$x = a\sin\theta$	$dx = a\cos\theta\, d\theta$	$\sqrt{a^2-x^2} = a\cos\theta$	$\theta = \sin^{-1}\left(\frac{x}{a}\right)$
$\sqrt{a^2+x^2}$ or a^2+x^2	$x = a\tan\theta$	$dx = a\sec^2\theta\, d\theta$	$\sqrt{a^2+x^2} = a\sec\theta$	$\theta = \tan^{-1}\left(\frac{x}{a}\right)$
$\sqrt{x^2-a^2}$	$x = a\sec\theta$	$dx = a\sec\theta\tan\theta\, d\theta$	$\sqrt{x^2-a^2} = a\tan\theta$	$\theta = \sec^{-1}\left(\frac{x}{a}\right)$

To emphasize the technique, we redo the computation for $\int \sqrt{1-x^2}\, dx$.

Example 2.27: Sine Substitution

Evaluate $\int \sqrt{1-x^2}\, dx$.

Solution. Since $\sqrt{1-x^2}$ appears in the integrand we try the trigonometric substitution $x = \sin\theta$. (Here we are using the restricted sine function with $\theta \in [-\pi/2,\ \pi/2]$ but typically omit this detail when writing out the solution.) Then $dx = \cos\theta\, d\theta$.

$$\begin{aligned}
\int \sqrt{1-x^2}\, dx &= \int \sqrt{1-\sin^2\theta}\cos\theta\, d\theta && \text{Using our (inverse) substitution}\\
&= \int \sqrt{\cos^2\theta}\cos\theta\, d\theta && \text{Since } \sin^2\theta+\cos^2\theta = 1\\
&= \int |\cos\theta|\cdot\cos\theta\, d\theta && \text{Since } \sqrt{\cos^2\theta} = |\cos\theta|\\
&= \int \cos^2\theta\, d\theta && \text{Since for } \theta\in[-\tfrac{\pi}{2},\tfrac{\pi}{2}] \text{ we have } \cos\theta \geq 0.
\end{aligned}$$

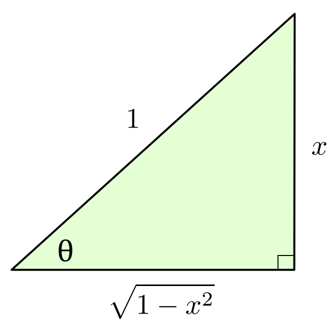
Often we omit the step containing the absolute value by our discussion above. Now, to integrate a power of cosine we use the guideline for products of sine and cosine and make use of the identity

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

Our integral then becomes

$$\int \sqrt{1-x^2}\, dx = \frac{1}{2}\int (1+\cos(2\theta))\, d\theta = \frac{\theta}{2} + \frac{\sin(2\theta)}{4} + C$$

To write the answer back in terms of x we use a right triangle. Since $\sin\theta = x/1$ we have the triangle:



The triangle gives $\sin\theta$, $\cos\theta$, $\tan\theta$, but have a $\sin(2\theta)$. Thus, we use an identity to write

$$\sin(2\theta) = 2\sin\theta\cos\theta = 2\left(\frac{x}{1}\right)\left(\frac{\sqrt{1-x^2}}{1}\right)$$

For θ by itself we use $\theta = \sin^{-1}x$. Thus, the integral is

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

♣

Example 2.28: Secant Substitution

Evaluate $\int \frac{\sqrt{25x^2-4}}{x}\, dx.$

Solution. We do not have $\sqrt{x^2-a^2}$ because of the 25, but if we factor 25 out we get:

$$\int \frac{\sqrt{25(x^2-(4/25))}}{x}\,dx = \int 5\frac{\sqrt{x^2-(4/25)}}{x}\,dx.$$

Now, $a=2/5$, so let $x=\frac{2}{5}\sec\theta$. Alternatively, we can think of the integral as being:

$$\int \frac{\sqrt{(5x)^2-4}}{x}\,dx$$

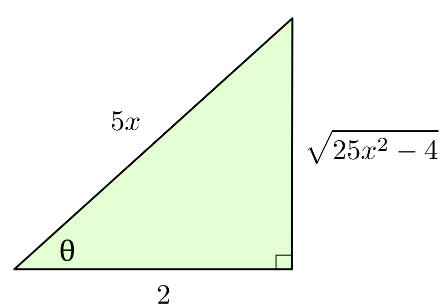
Then we could let $u=5x$ followed by $u=2\sec\theta$, etc. Or equivalently, we can avoid a u-substitution by letting $5x=2\sec\theta$. In either case we are using the trigonometric substitution $x=\frac{2}{5}\sec\theta$, but do use the method that makes the most sense to you! As $x=\frac{2}{5}\sec\theta$ we have $dx=\frac{2}{5}\sec\theta\tan\theta\,d\theta$.

$$\begin{aligned}
\int \frac{\sqrt{25x^2-4}}{x}\,dx &= \int \frac{\sqrt{25\frac{4\sec^2\theta}{25}-4}}{\frac{2}{5}\sec\theta}\frac{2}{5}\sec\theta\tan\theta\,d\theta && \text{Using the substitution}\\
&= \int \sqrt{4(\sec^2\theta-1)}\cdot\tan\theta\,d\theta && \text{Cancelling}\\
&= 2\int \sqrt{\tan^2\theta}\cdot\tan\theta\,d\theta && \text{Using } \tan^2\theta+1=\sec^2\theta\\
&= 2\int \tan^2\theta\,d\theta && \text{Simplifying}\\
&= 2\int (\sec^2\theta-1)\,d\theta && \text{Using } \tan^2\theta+1=\sec^2\theta\\
&= 2(\tan\theta-\theta)+C && \text{Since } \int\sec^2\theta\,d\theta=\tan\theta+C
\end{aligned}$$

For $\tan\theta$, we use a right triangle.

$$x=\frac{2}{5}\sec\theta \quad\to\quad x=\frac{2}{5}\frac{1}{\cos\theta} \quad\to\quad \cos\theta=\frac{2}{5x}$$

Using SOH CAH TOA, the triangle is then



For θ by itself, we use $\theta=\sec^{-1}(5x/2)$. Thus,

$$\int \frac{\sqrt{25x^2-4}}{x}\,dx = 2\left(\frac{\sqrt{25x^2-4}}{2}-\sec^{-1}\left(\frac{5x}{2}\right)\right)+C=\sqrt{25x^2-4}-2\sec^{-1}\left(\frac{5x}{2}\right)+C$$

♣

In the context of the previous example, some resources give an alternate guideline when choosing a trigonometric substitution.

$$\sqrt{a^2 - b^2x^2} \quad \to \quad x = \frac{a}{b}\sin\theta$$

$$\sqrt{b^2x^2 + a^2} \text{ or } (b^2x^2 + a^2) \quad \to \quad x = \frac{a}{b}\tan\theta$$

$$\sqrt{b^2x^2 - a^2} \quad \to \quad x = \frac{a}{b}\sec\theta$$

We next look at a tangent substitution.

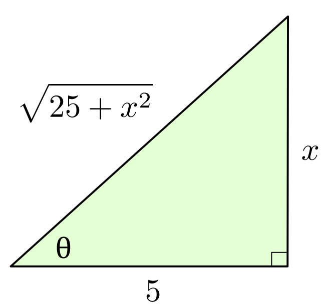
Example 2.29: Tangent Substitution

Evaluate $\displaystyle\int \frac{1}{\sqrt{25+x^2}}\,dx.$

Solution. Let $x = 5\tan\theta$ so that $dx = 5\sec^2\theta\,d\theta$.

$$\begin{aligned}
\int \frac{1}{\sqrt{25+x^2}}\,dx &= \int \frac{1}{\sqrt{25+25\tan^2\theta}}\,5\sec^2\theta\,d\theta && \text{Using our substitution}\\
&= \int \frac{1}{\sqrt{25(1+\tan^2\theta)}}\cdot 5\sec^2\theta\,d\theta && \text{Factor out 25}\\
&= \int \frac{1}{5\sqrt{sec^2\theta}}\cdot 5\sec^2\theta\,d\theta && \text{Using } \tan^2\theta + 1 = \sec^2\theta\\
&= \int \sec\theta\,d\theta && \text{Simplifying}\\
&= \ln|\sec\theta + \tan\theta| + C && \text{By } \int \sec\theta\,dx = \ln|\sec\theta+\tan\theta| + C
\end{aligned}$$

Since $\tan\theta = x/5$, we draw a triangle:



Then

$$\sec\theta = \frac{1}{\cos\theta} = \frac{\sqrt{25+x^2}}{5}.$$

Therefore, the integral is

$$\int \frac{1}{\sqrt{25+x^2}}\,dx = \ln\left|\frac{\sqrt{25+x^2}}{5} + \frac{x}{5}\right| + C$$

♣

In the next example, we will use the technique of completing the square in order to rewrite the integrand.

Example 2.30: Completing the Square

Evaluate $\displaystyle\int \frac{x}{\sqrt{3-2x-x^2}}\,dx$.

Solution. First, complete the square to write

$$3-2x-x^2 = 4-(x+1)^2$$

Now, we may let $u = x+1$ so that $du = dx$ (note that $x = u-1$) to get:

$$\int \frac{x}{\sqrt{4-(x+1)^2}}\,dx = \int \frac{u-1}{\sqrt{4-u^2}}\,du$$

Let $u = 2\sin\theta$ giving $du = 2\cos\theta\,d\theta$:

$$\int \frac{u-1}{\sqrt{4-u^2}}\,du = \int \frac{2\sin\theta-1}{2\cos\theta}\cdot 2\cos\theta\,d\theta = \int (2\sin\theta-1)\,d\theta$$

Integrating and using a triangle we get:

$$\begin{aligned}
\int \frac{x}{\sqrt{3-2x-x^2}} &= -2\cos\theta-\theta+C \\
&= -\sqrt{4-u^2}-\sin^{-1}\left(\frac{u}{2}\right)+C \\
&= -\sqrt{3-2x-x^2}-\sin^{-1}\left(\frac{x+1}{2}\right)+C
\end{aligned}$$

Note that in this problem we could have skipped the u-substitution if instead we let $x+1 = 2\sin\theta$. (For the triangle we would then use $\sin\theta = \dfrac{x+1}{2}$.) ♣

Exercises for Section 2.3

Exercise 2.3.1 *Evaluate the following indefinite integrals.*

(a) $\displaystyle\int \sqrt{x^2-1}\,dx$

(b) $\displaystyle\int \sqrt{9+4x^2}\,dx$

(c) $\displaystyle\int x\sqrt{1-x^2}\,dx$

(d) $\displaystyle\int x^2\sqrt{1-x^2}\,dx$

(e) $\displaystyle\int \frac{1}{\sqrt{1+x^2}}\,dx$

(f) $\displaystyle\int \sqrt{x^2+2x}\,dx$

(g) $\displaystyle\int \frac{1}{x^2(1+x^2)}\,dx$

(h) $\displaystyle\int \frac{x^2}{\sqrt{4-x^2}}\,dx$

(i) $\displaystyle\int \frac{\sqrt{x}}{\sqrt{1-x}}\,dx$ (j) $\displaystyle\int \frac{x^3}{\sqrt{4x^2-1}}\,dx$

2.4 Integration by Parts

We have already seen in Section 2.2.2 that recognizing the Product Rule can be useful, when we noticed that

$$\int \sec^3 u + \sec u \tan^2 u \, du = \sec u \tan u.$$

As with substitution, we do not have to rely on insight or cleverness to discover such antiderivatives; there is a technique that will often help to uncover the Product Rule called **Integration by Parts**.

Start with the Product Rule:

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

We can apply integration to this equation and obtain

$$f(x)g(x) = \int f'(x)g(x)\,dx + \int f(x)g'(x)\,dx,$$

and then rewrite this as

$$\int f(x)g'(x)\,dx = f(x)g(x) - \int f'(x)g(x)\,dx.$$

This may not seem particularly useful at first glance, but it turns out that in many cases we have an integral of the form

$$\int f(x)g'(x)\,dx$$

but that

$$\int f'(x)g(x)\,dx$$

is easier to integrate.

This technique for turning one integral into another is called **Integration by Parts**, and is usually written in more compact form.

Theorem 2.31: Integration by Parts

Let u and v be differentiable functions, then

$$\int u\,dv = uv - \int v\,du,$$

where

$$u = f(x) \quad \text{and} \quad v = g(x) \quad \text{so that} \quad du = f'(x)\,dx \quad \text{and} \quad dv = g'(x)\,dx.$$

Note:

1. To use this technique we need to identify likely candidates for $u = f(x)$ and $dv = g'(x)\,dx$. When choosing u and dv, keep in mind that we need to be able to readily find an antiderivative for dv and that du becomes simpler than u. Simpler could mean the power is reduced by one degree, or the original integral appears on the right side, or ...

2. After we have applied Integration by Parts, we then need to integrate $\int v\,du$. There is a danger to fall into a *circular trap* by choosing as the part to integrate (v) the term in the differential (du) from the first application of Integration by Parts. This does not provide you with any new information, but instead brings you back to the original integral. For example:

   To evaluate $\int x^2 \sin x\,dx$ we choose

   $$u = x^2 \quad \text{and} \quad v = \sin x\,dx \quad \text{so that} \quad du = 2x\,dx \quad \text{and} \quad dv = -\cos x\,dx,$$

   then

   $$\int x^2 \sin x\,dx = -x^2\cos x + 2\int x\cos x\,dx.$$

   If we ignore that the new integral is simpler than the original integral, which would tell us to continue in the same manner of selecting u and dv, we may fall into the circular trap of choosing

   $$u = \cos x \quad \text{and} \quad v = x\,dx \quad \text{so that} \quad du = -\sin x\,dx \quad \text{and} \quad dv = \frac{x^2}{2}\,dx,$$

   so that

   $$\begin{aligned}\int x^2 \sin x\,dx &= -x^2\cos x + 2\int x\cos x\,dx \\ &= -x^2\cos x + 2\left(\cos x\frac{x^2}{2} + \int \frac{x^2}{2}\sin x\,dx\right) \\ &= \int x^2 \sin x\,dx.\end{aligned}$$

   This shows that with our carelessness we have wasted our time and are back at the beginning.

Example 2.32: Product of a Linear Function and Logarithm

Evaluate $\int x\ln x\,dx$.

Solution. Let $u = \ln x$ so $du = 1/x\,dx$. Then we must let $dv = x\,dx$ so $v = x^2/2$ and

$$\int x\ln x\,dx = \frac{x^2\ln x}{2} - \int \frac{x^2}{2}\frac{1}{x}\,dx = \frac{x^2\ln x}{2} - \int \frac{x}{2}\,dx = \frac{x^2\ln x}{2} - \frac{x^2}{4} + C.$$

♣

Example 2.33: Inverse Trigonometric Function

Evaluate $\int \sin^{-1} x\,dx$.

Solution. Let $u = \sin^{-1} x$ so
$$du = \frac{1}{\sqrt{1-x^2}}\,dx.$$
Then we must let $dv = dx$ so $v = x$. Therefore,
$$\int \sin^{-1} x\,dx = x\sin^{-1} x - \int \frac{x}{\sqrt{1-x^2}}\,dx.$$
Now use substitution with $u = 1-x^2$ and $du = -2x\,dx$. Then
$$\begin{aligned}\int \frac{x}{\sqrt{1-x^2}}\,dx &= -\frac{1}{2}\int \frac{1}{\sqrt{u}}\,du \\ &= -\sqrt{u}+C \\ &= -\sqrt{1-x^2}+C.\end{aligned}$$
Altogether, we find that
$$\int \sin^{-1} x\,dx = x\sin^{-1} x + \sqrt{1-x^2}+C.$$

♣

Example 2.34: Secant Cubed (again)

Evaluate $\int \sec^3 x\,dx$.

Solution. Of course we already know the answer to this, but we needed to be clever to discover it. Here we'll use the new technique to discover the antiderivative. Let $u = \sec x$ and $dv = \sec^2 x\,dx$. Then $du = \sec x \tan x$ and $v = \tan x$ and
$$\begin{aligned}\int \sec^3 x\,dx &= \sec x\tan x - \int \tan^2 x \sec x\,dx \\ &= \sec x\tan x - \int (\sec^2 x - 1)\sec x\,dx \\ &= \sec x\tan x - \int \sec^3 x\,dx + \int \sec x\,dx.\end{aligned}$$

At first this looks useless—we're right back to $\int \sec^3 x\,dx$. But looking more closely, we notice that we can add this integral to both sides and are left to deal with the integral $\int \sec x\,dx$.
$$\int \sec^3 x\,dx = \sec x\tan x - \int \sec^3 x\,dx + \int \sec x\,dx$$

$$\begin{aligned}
\int \sec^3 x\,dx + \int \sec^3 x\,dx &= \sec x \tan x + \int \sec x\,dx \\
2\int \sec^3 x\,dx &= \sec x \tan x + \int \sec x\,dx \\
\int \sec^3 x\,dx &= \frac{\sec x \tan x}{2} + \frac{1}{2}\int \sec x\,dx
\end{aligned}$$

♣

Now we use our knowledge of $\int \sec x\,dx$ to conclude that

$$\int \sec^3 x\,dx = \frac{\sec x \tan x}{2} + \frac{\ln|\sec x + \tan x|}{2} + C.$$

Example 2.35: Product of a Polynomial and Trigonometric Function

Evaluate $\int x^2 \sin x\,dx$.

Solution. Let $u = x^2$, $dv = \sin x\,dx$; then $du = 2x\,dx$ and $v = -\cos x$. Now

$$\int x^2 \sin x\,dx = -x^2 \cos x + \int 2x \cos x\,dx.$$

This is better than the original integral since the power of x has been reduced by one degree, but we need to do Integration by Parts again. Let $u = 2x$, $dv = \cos x\,dx$; then $du = 2$ and $v = \sin x$, and

$$\begin{aligned}
\int x^2 \sin x\,dx &= -x^2 \cos x + \int 2x \cos x\,dx \\
&= -x^2 \cos x + 2x \sin x - \int 2 \sin x\,dx \\
&= -x^2 \cos x + 2x \sin x + 2\cos x + C.
\end{aligned}$$

♣

2.4.1. Tabular Method

Such repeated use of Integration by Parts is fairly common, but it can be a bit tedious to accomplish, and it is easy to make errors, especially sign errors involving the subtraction in the formula. There is a nice **tabular method** to accomplish the calculation that minimizes the chance for error and speeds up the whole process. We illustrate with the previous example. Here is the table:

sign	u	dv
+	x^2	$\sin x$
-	$2x$	$-\cos x$
+	2	$-\sin x$
-	0	$\cos x$

To form this table, we start with u at the top of the second column and repeatedly compute the derivative; starting with dv at the top of the third column, we repeatedly compute the antiderivative. In the first column, we place a "$-$" in every second row. To form the second table we combine the first and second columns by ignoring the boundary; if you do this by hand, you may simply start with two columns and add a "$-$" to every second row.

Alternatively, we can use the following table:

u	**dv**
x^2	$\sin x$
$-2x$	$-\cos x$
2	$-\sin x$
0	$\cos x$

To compute with this second table we begin at the top. Multiply the first entry in column u by the second entry in column dv to get $-x^2\cos x$, and add this to the integral of the product of the second entry in column u and second entry in column dv. This gives:

$$-x^2\cos x + \int 2x\cos x\,dx,$$

or exactly the result of the first application of Integration by Parts. Since this integral is not yet easy, we return to the table. Now we multiply twice on the diagonal, $(x^2)(-\cos x)$ and $(-2x)(-\sin x)$ and then once straight across, $(2)(-\sin x)$, and combine these as

$$-x^2\cos x + 2x\sin x - \int 2\sin x\,dx,$$

giving the same result as the second application of Integration by Parts. While this integral is easy, we may return yet once more to the table. Now multiply three times on the diagonal to get $(x^2)(-\cos x)$, $(-2x)(-\sin x)$, and $(2)(\cos x)$, and once straight across, $(0)(\cos x)$. We combine these as before to get

$$-x^2\cos x + 2x\sin x + 2\cos x + \int 0\,dx = -x^2\cos x + 2x\sin x + 2\cos x + C.$$

Typically we would fill in the table one line at a time, until the "straight across" multiplication gives an easy integral. If we can see that the u column will eventually become zero, we can instead fill in the whole table; computing the products as indicated will then give the entire integral, including the "$+C$", as above.

Exercises for Section 2.4

Exercise 2.4.1 *Evaluate the following indefinite integrals. Hint: The original integral will appear in your solution process.*

(a) $\int xe^{x^2}\,dx$

(b) $\int \sin^2 x\,dx$

(c) $\int x \arctan x\,dx$

(d) $\int x\sin^2 x\,dx$

(e) $\int x\sin x\cos x\,dx$

Exercise 2.4.2 *Evaluate the following indefinite integrals.*

(a) $\int x\cos x\,dx$

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

(c) $\int xe^x\,dx$

(d) $\int \ln x\,dx$

(e) $\int x^3\sin x\,dx$

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

(g) $\int \arctan(\sqrt{x})\,dx$

(h) $\int \sin(\sqrt{x})\,dx$

(i) $\int \sec^2 x\csc^2 x\,dx$

2.5 Partial Fraction Method for Rational Functions

A **rational function** is a fraction with polynomials in the numerator and denominator. For example,

$$\frac{x^3}{x^2+x-6}, \qquad \frac{1}{(x-3)^2}, \qquad \frac{x^2+1}{x^2-1},$$

are all rational functions of x. There is a general technique called the **Partial Fraction Method** that, in principle, allows us to integrate any rational function. The algebraic steps in the technique are rather cumbersome if the polynomial in the denominator has degree more than 2, and the technique requires that we factor the denominator, something that is not always possible. However, in practice one does not often run across rational functions with high degree polynomials in the denominator for which one has to find the antiderivative function. So we shall explain how to find the antiderivative of a *reduced* rational function only when the denominator is a quadratic polynomial ax^2+bx+c.

2.5.1. Using Substitution Rule with Rational Fractions

We should mention a special type of rational function that we already know how to integrate: If the denominator has the form $(ax+b)^n$, the substitution $u = ax+b$ will always work. The denominator becomes u^n, and each x in the numerator is replaced by $(u-b)/a$ and $dx = du/a$. While it may be tedious to complete the integration if the numerator has high degree, it is merely a matter of algebra.

Example 2.36: Substitution and Splitting Up a Fraction

Find $\int \frac{x^3}{(3-2x)^5}\,dx.$

Solution. Using the substitution $u = 3 - 2x$ we obtain

$$du = -2dx, \quad \text{and} \quad -\frac{u-3}{2} = x$$

so that

$$\int \frac{x^3}{(3-2x)^5}\,dx \;=\; \frac{1}{-2}\int \frac{\left(\frac{u-3}{-2}\right)^3}{u^5}\,du$$

We now divide through by the simple denominator to obtain powers of u:

$$\begin{aligned}
\int \frac{x^3}{(3-2x)^5}\,dx &= \frac{1}{16}\int \frac{u^3 - 9u^2 + 27u - 27}{u^5}\,du \\
&= \frac{1}{16}\int u^{-2} - 9u^{-3} + 27u^{-4} - 27u^{-5}\,du \\
&= \frac{1}{16}\left(\frac{u^{-1}}{-1} - \frac{9u^{-2}}{-2} + \frac{27u^{-3}}{-3} - \frac{27u^{-4}}{-4}\right) + C \\
&= \frac{1}{16}\left(-u^{-1} + \frac{9u^{-2}}{2} - 9u^{-3} + \frac{27u^{-4}}{4}\right) + C
\end{aligned}$$

All that is left to do is replace u with our substitution.

$$\begin{aligned}
\int \frac{x^3}{(3-2x)^5}\,dx &= \frac{1}{16}\left(-(3-2x)^{-1} + \frac{9(3-2x)^{-2}}{2} - 9(3-2x)^{-3} + \frac{27(3-2x)^{-4}}{4}\right) + C \\
&= -\frac{1}{16(3-2x)} + \frac{9}{32(3-2x)^2} - \frac{9}{16(3-2x)^3} + \frac{27}{64(3-2x)^4} + C
\end{aligned}$$

♣

2.5.2. Denominator with Distinct Linear Factors

We now proceed to the case in which the denominator is a quadratic polynomial. We can always factor out the coefficient of x^2 and put it outside the integral, so we can assume that the denominator has the form $x^2 + bx + c$. There are three possible cases, depending on how the quadratic factors: either $x^2 + bx + c = (x-r)(x-s)$, $x^2 + bx + c = (x-r)^2$, or the quadratic factor is irreducible, i.e. it doesn't factor. We can use the quadratic formula to decide which of these we have, and to factor the quadratic if it is possible.

If $x^2 + bx + c = (x-r)^2$ then we have the special case we have already seen, that can be handled with a substitution.

If $x^2 + bx + c = (x-r)(x-s)$, we have an integral of the form

$$\int \frac{p(x)}{(x-r)(x-s)}\,dx$$

where $p(x)$ is a polynomial. The first step is to make sure that $p(x)$ has degree less than 2.

Example 2.37: Review of Long Division

Compute $\dfrac{x^3-2x^2-4}{x-3}$, *that is, divide* $x-3$ *into* x^3-2x^2-4.

Solution. We first write the dividend as x^3-2x^2+0x-4 (that is, include 0's). Now, divide the first term of the dividend by the highest term of the divisor:

$$
\begin{array}{rl}
 & x^2 \\
x-3 & \overline{)\,x^3-2x^2+0x-4}
\end{array}
\qquad \text{Here } x^3 \div x = x^2
$$

Next, we multiply the divisor by this result and write it below:

$$
\begin{array}{rl}
 & x^2 \\
x-3 & \overline{)\,x^3-2x^2+0x-4} \\
 & \;\; x^3-3x^2
\end{array}
\qquad \text{Here } x^2\cdot(x-3) = x^3-3x^2
$$

Subtract and write the result below:

$$
\begin{array}{rl}
 & x^2 \\
x-3 & \overline{)\,x^3-2x^2+0x-4} \\
 & \underline{\;\; x^3-3x^2\;\;} \\
 & \qquad +x^2+0x-4
\end{array}
$$

We repeat the above steps of dividing, multiplying then subtracting:

$$
\begin{array}{rl}
 & x^2\;+\;x \\
x-3 & \overline{)\,x^3-2x^2+0x-4} \\
 & \underline{\;\; x^3-3x^2\;\;} \\
 & \qquad +x^2+0x-4
\end{array}
\;\rightarrow\;
\begin{array}{rl}
 & x^2\;+\;x \\
x-3 & \overline{)\,x^3-2x^2+0x-4} \\
 & \underline{\;\; x^3-3x^2\;\;} \\
 & \qquad +x^2+0x-4 \\
 & \qquad +x^2-3x
\end{array}
\;\rightarrow\;
\begin{array}{rl}
 & x^2\;+\;x \\
x-3 & \overline{)\,x^3-2x^2+0x-4} \\
 & \underline{\;\; x^3-3x^2\;\;} \\
 & \qquad +x^2+0x-4 \\
 & \qquad \underline{+x^2-3x} \\
 & \qquad\qquad\quad 3x-4
\end{array}
$$

We divide:

$$
\begin{array}{rl}
 & x^2\;+\;x\;+\;3 \\
x-3 & \overline{)\,x^3-2x^2+0x-4} \\
 & \underline{\;\; x^3-3x^2\;\;} \\
 & \qquad +x^2+0x-4 \\
 & \qquad \underline{+x^2-3x} \\
 & \qquad\qquad\quad 3x-4
\end{array}
$$

Then multiply and subtract:

$$
\begin{array}{rl}
 & x^2\;+\;x\;+\;3 \\
x-3 & \overline{)\,x^3-2x^2+0x-4} \\
 & \underline{\;\; x^3-3x^2\;\;} \\
 & \qquad +x^2+0x-4 \\
 & \qquad \underline{+x^2-3x} \\
 & \qquad\qquad\quad 3x-4 \\
 & \qquad\qquad\quad 3x-9
\end{array}
\;\rightarrow\;
\begin{array}{rl}
 & x^2\;+\;x\;+\;3 \\
x-3 & \overline{)\,x^3-2x^2+0x-4} \\
 & \underline{\;\; x^3-3x^2\;\;} \\
 & \qquad +x^2+0x-4 \\
 & \qquad \underline{+x^2-3x} \\
 & \qquad\qquad\quad 3x-4 \\
 & \qquad\qquad\quad \underline{3x-9} \\
 & \qquad\qquad\qquad\;\; 5
\end{array}
$$

The *quotient* is x^2+x+3 and the *remainder* is 5. Therefore,

$$\frac{x^3-2x^2-4}{x-3}=(x^2+x+3)+\frac{5}{x-3}$$

This means that $(x-3)$ divides (x^3-2x^2-4) a total of (x^2+x+3) times with 5 leftover.

♣

Now consider the following simple algebra of fractions:

$$\frac{A}{x-r}+\frac{B}{x-s}=\frac{A(x-s)+B(x-r)}{(x-r)(x-s)}=\frac{(A+B)x-As-Br}{(x-r)(x-s)}.$$

That is, adding two fractions with constant numerator and denominators $(x-r)$ and $(x-s)$ produces a fraction with denominator $(x-r)(x-s)$ and a polynomial of degree less than 2 for the numerator. We want to reverse this process: Starting with a single fraction, we want to write it as a sum of two simpler fractions. An example should make it clear how to proceed.

Example 2.38: Distinct Linear Factors

Compute $\displaystyle\int \frac{-1}{x^2-2x-3}\,dx.$

Solution. We first notice that the degree of the numerator is less than the degree of the denominator. Next, we factor the denominator:

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

Therefore, the partial fraction decomposition of the integrand is

$$\frac{-1}{(x+1)(x+3)}=\frac{A}{x+1}+\frac{B}{x-3}=\frac{A(x-3)+B(x+1)}{(x+1)(x+3)}.$$

We now present two methods for determining the constants A and B.
Method 1: We notice that we can plug in values for x that give zero somewhere. If we take $x=3$, we get

$$-1=A(3-3)=B(3+1)\implies B=-\frac{1}{4},$$

and that if we take $x=-1$, we get

$$-1=A(-1-3)+B(-1+1)\implies A=\frac{1}{4}.$$

Method 2: We expand the right hand side,

$$0\cdot x-1=(A+B)x+(-3A+B)$$

and compare the polynomial coefficients: By comparing the x^1 terms, we see that we require $0=A+B$, and comparing the x^0 terms gives $-1=-3A=B$. We now solve the resulting system of equations.

$$\begin{array}{c}0=A+B\\-1=-3A+B\end{array}\implies\begin{array}{c}A=-B\\-1=-3A+B\end{array}\implies\begin{array}{c}A=-B\\-1=4B\end{array}\implies\begin{array}{c}A=\frac{1}{4}\\B=-\frac{1}{4}\end{array}$$

Thus, the partial fraction decomposition of the integrand is

$$\int \frac{-1}{x^2-2x-3}\,dx = \int \frac{\frac{1}{4}}{x+1} + \frac{-\frac{1}{4}}{x-3}\,dx,$$

and we can now integrate:

$$\begin{aligned}\int \frac{-1}{x^2-2x-3}\,dx &= \int \frac{1/4}{x+1} + \frac{-1/4}{x-3}\,dx \\ &= \frac{1}{4}\int \frac{1}{x+1}\,dx - \frac{1}{4}\int \frac{1}{x-3}\,dx \\ &= \frac{1}{4}\ln|x+1| - \frac{1}{4}\ln|x-3| + C.\end{aligned}$$

♣

Example 2.39: Partial Fraction Decomposition

Given $\displaystyle\int \frac{x^3}{(x-2)(x+3)}\,dx$

(a) Perform long division.

(b) Evaluate the integral.

Solution.

(a) We rewrite the integrand so that the numerator that has degree less than 2 using polynomial long division:

$$\frac{x^3}{(x-2)(x+3)} = \frac{x^3}{x^2+x-6} = x-1+\frac{7x-6}{x^2+x-6} = x-1+\frac{7x-6}{(x-2)(x+3)}.$$

Then

$$\int \frac{x^3}{(x-2)(x+3)}\,dx = \int (x-1)\,dx + \int \frac{7x-6}{(x-2)(x+3)}\,dx.$$

(b) The first integral is easy, so only the second requires some work. We start by writing $\dfrac{7x-6}{(x-2)(x+3)}$ as the sum of two fractions. We want to end up with

$$\frac{7x-6}{(x-2)(x+3)} = \frac{A}{x-2} + \frac{B}{x+3}.$$

If we go ahead and add the fractions on the right hand side, we seek a common denominator, and get:

$$\frac{7x-6}{(x-2)(x+3)} = \frac{(A+B)x+3A-2B}{(x-2)(x+3)}.$$

So all we need to do is find A and B so that $7x-6=(A+B)x+3A-2B$, which is to say, we need

$$\begin{cases} 7 & = A+B \\ -6 & = 3A-2B \end{cases}$$

This is a problem you've seen before: Solve a system of two equations in two unknowns. There are many ways to proceed; here's one: If $7=A+B$ then $B=7-A$ and so

$$-6 = 3A-2B = 3A-2(7-A) = 3A-14+2A = 5A-14.$$

This is easy to solve for A:

$$A = 8/5 \implies B = 7-A = 7-8/5 = 27/5.$$

Thus

$$\int \frac{7x-6}{(x-2)(x+3)}\,dx = \int \left(\frac{8}{5}\frac{1}{x-2} + \frac{27}{5}\frac{1}{x+3}\right)dx = \frac{8}{5}\ln|x-2| + \frac{27}{5}\ln|x+3| + C.$$

The answer to the original problem is now

$$\begin{aligned} \int \frac{x^3}{(x-2)(x+3)}\,dx &= \int (x-1)\,dx + \int \frac{7x-6}{(x-2)(x+3)}\,dx \\ &= \frac{x^2}{2} - x + \frac{8}{5}\ln|x-2| + \frac{27}{5}\ln|x+3| + C. \end{aligned}$$

♣

Example 2.40: Two Linear Factors in Disguise

Evaluate $\displaystyle\int \frac{\cos x}{4-\sin^2 x}\,dx.$

Solution. We start by making the substitution

$$w = \sin x \implies dw = \cos x\,dx.$$

Then,

$$\begin{aligned} \int \frac{\cos x}{4-\sin^2 x}\,dx &= \int \frac{dw}{4-w^2} \\ &= \int \frac{dw}{(2-w)(2+w)} \end{aligned}$$

We are now in a position to use the method of partial fraction decomposition. Solving

$$\frac{1}{(2-w)(2+w)} = \frac{A}{2-w} + \frac{B}{2+w} = \frac{w(A-B)+2(A+B)}{(2+w)(2-w)},$$

we require

$$\begin{cases} 0 & = A - B \\ 1 & = 2(A+B) \end{cases} \implies A = B = \frac{1}{4}.$$

Therefore,

$$\int \frac{dw}{(2-w)(2+w)} = \int \left(\frac{1/4}{2-w} + \frac{1/4}{2+w} \right) dw,$$

which is readily evaluated:

$$\int \left(\frac{1/4}{2-w} + \frac{1/4}{2+w} \right) dw = -\frac{1}{4} \ln|2-w| + \frac{1}{4}|2+w| + C$$

It remains to rewrite this expression in terms of x:

$$\int \frac{\cos x}{4-\sin^2 x}\, dx = \frac{1}{4} \ln \left| \frac{2+\sin x}{2-\sin x} \right| + C.$$

♣

2.5.3. Denominator with Irreducible Quadratic Factor

Now suppose that x^2+bx+c doesn't factor. Again we can use long division to ensure that the numerator has degree less than 2, then we complete the square.

Example 2.41: Denominator Does Not Factor

Evaluate $\displaystyle\int \frac{x+1}{x^2+4x+8}\, dx.$

Solution. The quadratic denominator does not factor. We could complete the square and use a trigonometric substitution, but it is simpler to rearrange the integrand:

$$\int \frac{x+1}{x^2+4x+8}\, dx = \int \frac{x+2}{x^2+4x+8}\, dx - \int \frac{1}{x^2+4x+8}\, dx.$$

The first integral is an easy substitution problem, using $u = x^2+4x+8$:

$$\int \frac{x+2}{x^2+4x+8}\, dx = \frac{1}{2} \int \frac{du}{u} = \frac{1}{2} \ln|x^2+4x+8|.$$

For the second integral we complete the square:

$$x^2+4x+8 = (x+2)^2+4 = 4\left(\left(\frac{x+2}{2} \right)^2 + 1 \right),$$

making the integral

$$\frac{1}{4} \int \frac{1}{\left(\frac{x+2}{2}\right)^2+1}\, dx.$$

Using $u = \dfrac{x+2}{2}$ we get

$$\frac{1}{4}\int \frac{1}{\left(\frac{x+2}{2}\right)^2+1}\,dx = \frac{1}{4}\int \frac{2}{u^2+1}\,du = \frac{1}{2}\arctan\left(\frac{x+2}{2}\right).$$

The final answer is now

$$\int \frac{x+1}{x^2+4x+8}\,dx = \frac{1}{2}\ln|x^2+4x+8| - \frac{1}{2}\arctan\left(\frac{x+2}{2}\right) + C.$$

♣

2.5.4. Summary

Many combinations of linear and quadratic factors are possible in the denominator of a rational function. However, we are not concerned with supplying the techniques needed to solve these types of rational functions. Their solution is simply a matter of algebraic skill in determining a partial fraction decomposition, which allows for the ready integration of each partial fraction. However, we want to conclude with the following emphasis and one example to give you a taste of different combinations of factors than shown in this section.

Note:

1. Unless we are able to factor the denominator, we are unable to use the Partial Fraction Method.
2. When considering the Partial Fraction Method, we must ensure that the degree of the numerator is smaller than that of the denominator. So brush up on your long division skill.

The following example alludes to the technique used for the partial fraction decomposition when there is an irreducible quadratic factor.

Example 2.42: Cubic Denominator

Evaluate $\displaystyle\int \frac{x^2-2x+1}{x^3+x}\,dx.$

Solution. We first factor the denominator,

$$\int \frac{x^2-2x+1}{x^3+x}\,dx = \int \frac{x^2-2x+1}{x(x^2+1)}\,dx.$$

Now, let

$$\frac{x^2-2x+1}{x(x^2+1)} = \frac{A}{x} + \frac{Bx+C}{x^2+1} = \frac{Ax^2+A+Bx^2+Cx}{x(x^2+1)}.$$

Matching terms gives,

$$\begin{cases} 1 & = A+B \\ -2 & = C \\ 1 & = A \end{cases}$$

From the above system of equations, we see that we need $A = 1$, $C = -2$, and $B = 0$. Thus,

$$\begin{aligned}\int \frac{x^2-2x+1}{x^3+x}\,dx &= \int \left(\frac{1}{x} - \frac{2}{x^2+1}\right) dx \\ &= \ln|x| - 2\arctan x + C.\end{aligned}$$

♣

Exercises for Section 2.5

Exercise 2.5.1 *Evaluate the following indefinite integrals.*

(a) $\displaystyle\int \frac{1}{4-x^2}\,dx$

(b) $\displaystyle\int \frac{x^4}{4-x^2}\,dx$

(c) $\displaystyle\int \frac{1}{x^2+10x+25}\,dx$

(d) $\displaystyle\int \frac{x^2}{4-x^2}\,dx$

(e) $\displaystyle\int \frac{x^4}{4+x^2}\,dx$

(f) $\displaystyle\int \frac{1}{x^2+10x+29}\,dx$

(g) $\displaystyle\int \frac{x^3}{4+x^2}\,dx$

(h) $\displaystyle\int \frac{1}{x^2+10x+21}\,dx$

(i) $\displaystyle\int \frac{1}{2x^2-x-3}\,dx$

(j) $\displaystyle\int \frac{1}{x^2+3x}\,dx$

2.6 Numerical Integration

We have now seen some of the most generally useful methods for discovering antiderivatives, and there are others. Unfortunately, some functions have no simple antiderivatives. In such cases, if the value of a definite integral is needed it will have to be approximated.

2.6.1. Midpoint Rule

Of course, we already know from Section 1.3 one way to approximate an integral: If we think of the integral as computing an area, we can add up the areas of some rectangles (Riemann sum). While this is quite simple, it is usually the case that a large number of rectangles is needed to get acceptable accuracy. As pointed out before, the Midpoint Rule for the Riemann sum works best in most approximations that are based on rectangles.

Theorem 2.43: The Midpoint Rule

Let $f(x)$ be defined on a closed interval $[a,b]$ that is subdivided into n subintervals of equal length $\Delta x = (b-a)/n$ using $n+1$ points $x_i = a + i\Delta x$:

$$x_0 = a, x_1 = a + \Delta x, \ldots, x_{n-1} = a + (n-1)\Delta x, x_n = b.$$

Then the integral $\int_a^b f(x)\,dx$ can be approximated by

$$\int_a^b f(x)\,dx \approx \sum_{i=1}^{n} f\left(\frac{x_{i-1}+x_i}{2}\right)\Delta x.$$

In practice, an approximation is useful only if we know how accurate it is; for example, we might need a particular value accurate to three decimal places. When we compute a particular approximation to an integral, the error is the difference between the approximation and the true value of the integral. For any approximation technique, we need an **error bound**, a value that is guaranteed to be larger than the actual error. If A is an approximation and E is the associated error bound, then we know that the true value of the integral is between $A-E$ and $A+E$. In the case of our approximation of the integral, we want $E = E(\Delta x)$ to be a function of Δx that gets small rapidly as Δx gets small. Fortunately, for many functions, there is such an error bound associated with the midpoint approximation.

Theorem 2.44: Error for Midpoint Approximation

Suppose f has a second derivative f'' everywhere on the interval $[a,b]$, and $|f''(x)| \le M$ for all x in the interval. With $\Delta x = (b-a)/n$, an error bound for the midpoint approximation is

$$E(\Delta x) = \frac{b-a}{24} M (\Delta x)^2 = \frac{(b-a)^3}{24n^2} M.$$

We will see two other methods that work reasonably well and yet are fairly simple; in some cases more sophisticated techniques will be needed.

2.6.2. Trapezoid Rule

A similar approach is much better. We approximate the area under a curve over a small interval as the area of a trapezoid. This technique for approximating an integral is known as the **Trapezoid Rule**. In Figure 2.1 we see an area under a curve approximated by rectangles and by trapezoids; it is apparent that the trapezoids give a substantially better approximation on each subinterval.

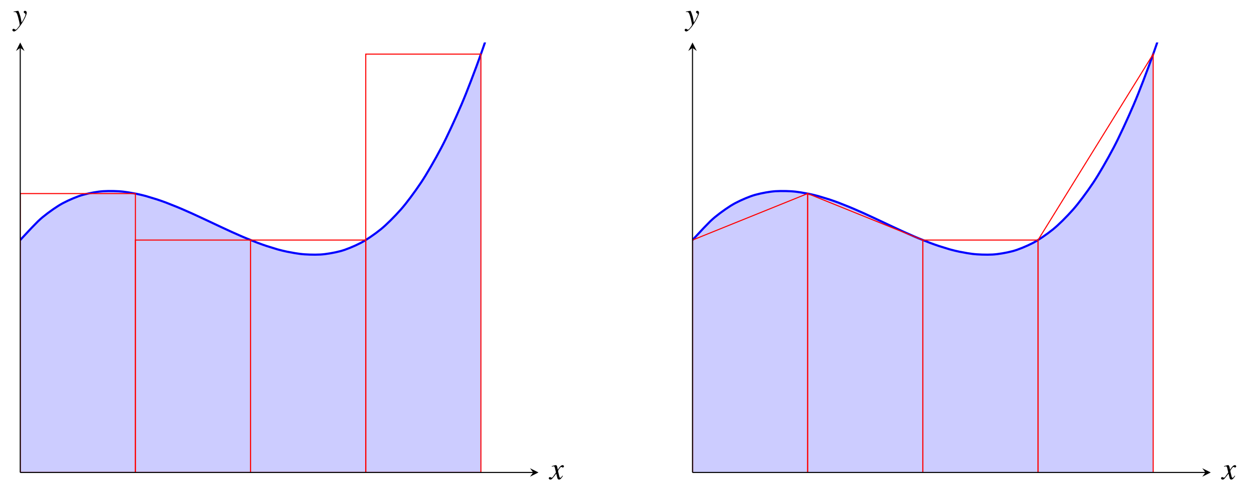


Figure 2.1: Approximating an area with rectangles and with trapezoids.

As with rectangles, we divide the interval into n equal subintervals of length Δx. A typical trapezoid is pictured in Figure 2.2; it has area

$$\frac{f(x_i)+f(x_{i+1})}{2}\Delta x.$$

If we add up the areas of all trapezoids we get

$$\begin{aligned}
&\frac{f(x_0)+f(x_1)}{2}\Delta x+\frac{f(x_1)+f(x_2)}{2}\Delta x+\cdots+\frac{f(x_{n-1})+f(x_n)}{2}\Delta x\\
&=\left(\frac{f(x_0)}{2}+f(x_1)+f(x_2)+\cdots+f(x_{n-1})+\frac{f(x_n)}{2}\right)\Delta x\\
&=\frac{\Delta x}{2}\left(f(x_0)+2f(x_1)+2f(x_2)+\cdots+2f(x_{n-1})+f(x_n)\right).
\end{aligned}$$

For a modest number of subintervals this is not too difficult to do with a calculator; a computer can easily handle many subintervals.

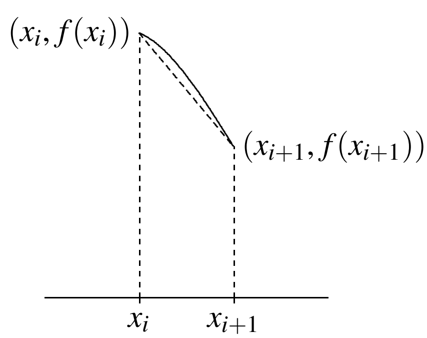


Figure 2.2: A single trapezoid.

We summarize this result in the theorem below.

Theorem 2.45: Trapezoid Rule

Let $f(x)$ be defined on a closed interval $[a,b]$ that is subdivided into n subintervals of equal length $\Delta x = (b-a)/n$ using $n+1$ points $x_i = a + i\Delta x$:

$$x_0 = a, x_1 = a + \Delta x, \ldots, x_{n-1} = a + (n-1)\Delta x, x_n = b.$$

Then the integral $\int_a^b f(x)\,dx$ can be approximated by

$$\int_a^b f(x)\,dx \approx \frac{\Delta x}{2}\left(f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n)\right).$$

As with the midpoint method, this is useful only with an **error bound**:

Theorem 2.46: Error for Trapezoid Approximation

Suppose f has a second derivative f'' everywhere on the interval $[a,b]$, and $|f''(x)| \le M$ for all x in the interval. With $\Delta x = (b-a)/n$, an error bound for the trapezoid approximation is

$$E(\Delta x) = \frac{b-a}{12} M (\Delta x)^2 = \frac{(b-a)^3}{12n^2} M.$$

Let's see how we can use this.

Example 2.47: Approximate an Integral With Trapezoids

Approximate $\int_0^1 e^{-x^2}\,dx$ to two decimal places.

Solution. The second derivative of $f = e^{-x^2}$ is $(4x^2-2)e^{-x^2}$, and it is not hard to see that on $[0,1]$ $|f''(x)|$ has a maximum value of 2, thus we begin by estimating the number of subintervals we are likely to need. To get two decimal places of accuracy, we will certainly need $E(\Delta x) < 0.005$ or

$$\begin{aligned} \frac{1}{12}(2)\frac{1}{n^2} &< 0.005 \\ \frac{1}{6}(200) &< n^2 \\ 5.77 \approx \sqrt{\frac{100}{3}} &< n \end{aligned}$$

With $n = 6$, the error bound is thus $1/6^3 < 0.0047$. We compute the trapezoid approximation for six intervals:

$$\left(\frac{f(0)}{2} + f(1/6) + f(2/6) + \cdots + f(5/6) + \frac{f(1)}{2}\right)\frac{1}{6} \approx 0.74512.$$

The error bound gives an estimate of the error on either side of the approximation, and so the true value of the integral is between $0.74512 - 0.0047 = 0.74042$ and $0.74512 + 0.0047 = 0.74982$. Unfortunately, the first rounds to 0.74 and the second rounds to 0.75, so we can't be sure of the correct value in the second decimal place; we need to pick a larger n. As it turns out, we need to go to $n = 12$ to get two bounds that both round to the same value, which turns out to be 0.75. For comparison, using 12 rectangles to approximate the area gives 0.7727, which is considerably less accurate than the approximation using six trapezoids.

In practice it generally pays to start by requiring better than the maximum possible error; for example, we might have initially required $E(\Delta x) < 0.001$, or

$$\begin{aligned} \frac{1}{12}(2)\frac{1}{n^2} &< 0.001 \\ \frac{1}{6}(1000) &< n^2 \\ 12.91 \approx \sqrt{\frac{500}{3}} &< n \end{aligned}$$

Had we immediately tried $n = 13$ this would have given us the desired answer. ♣

2.6.3. Simpson's Rule

The trapezoid approximation works well, especially compared to rectangles, because the tops of the trapezoids form a reasonably good approximation to the curve when Δx is fairly small. What if we try to approximate the curve more closely by using something other than a straight line in our search for a better approximation to the integral of f? The obvious candidate is a parabola as shown in Figure 2.3: If we can approximate a short piece of the curve with a parabola with equation $y = ax^2 + bx + c$, we can easily compute the area under the parabola.

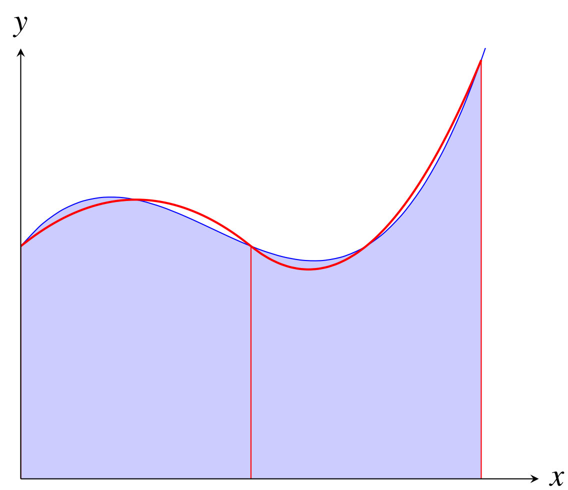


Figure 2.3: Approximating an area with parabolas.

There are an infinite number of parabolas through any two given points, but only one through three given points. If we find a parabola through three consecutive points $(x_i, f(x_i))$, $(x_{i+1}, f(x_{i+1}))$, $(x_{i+2}, f(x_{i+2}))$

on the curve, it should be quite close to the curve over the whole interval $[x_i, x_{i+2}]$, as in Figure 2.4. If we divide the interval $[a,b]$ into an even number of subintervals, we can then approximate the curve by a sequence of parabolas, each covering two of the subintervals. For this to be practical, we would like a simple formula for the area under one parabola, namely, the parabola through $(x_i, f(x_i))$, $(x_{i+1}, f(x_{i+1}))$, and $(x_{i+2}, f(x_{i+2}))$. That is, we should attempt to write down the parabola $y = ax^2 + bx + c$ through these points and then integrate it, and hope that the result is fairly simple. Although the algebra involved is messy, this turns out to be possible. The algebra is well within the capability of a good computer algebra system like Sage, so we will present the result without all of the algebra.

To find the parabola, we solve these three equations for a, b, and c:

$$\begin{aligned} f(x_i) &= a(x_{i+1} - \Delta x)^2 + b(x_{i+1} - \Delta x) + c \\ f(x_{i+1}) &= a(x_{i+1})^2 + b(x_{i+1}) + c \\ f(x_{i+2}) &= a(x_{i+1} + \Delta x)^2 + b(x_{i+1} + \Delta x) + c \end{aligned}$$

Not surprisingly, the solutions turn out to be quite messy. Nevertheless, Sage can easily compute and simplify the integral to get

$$\int_{x_{i+1}-\Delta x}^{x_{i+1}+\Delta x} ax^2 + bx + c\,dx = \frac{\Delta x}{3}(f(x_i) + 4f(x_{i+1}) + f(x_{i+2})).$$

Now the sum of the areas under all parabolas is

$$\begin{aligned} &\frac{\Delta x}{3}(f(x_0) + 4f(x_1) + f(x_2) + f(x_2) + 4f(x_3) + f(x_4) + \cdots + f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)) = \\ &\quad \frac{\Delta x}{3}(f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \cdots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)). \end{aligned}$$

This is just slightly more complicated than the formula for trapezoids; we need to remember the alternating 2 and 4 coefficients, and that the interval must be divided into an *even* number of subintervals. This approximation technique is referred to as **Simpson's Rule**.

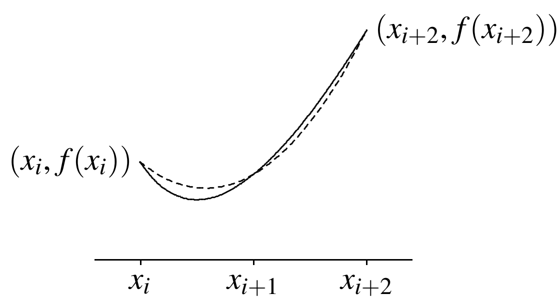


Figure 2.4: A parabola (dashed) approximating a curve (solid).

We capture our findings in the following theorem.

Theorem 2.48: Simpson's Rule

Let $f(x)$ be defined on a closed interval $[a,b]$ that is subdivided into n subintervals of equal length $\Delta x = (b-a)/n$ using $n+1$ points $x_i = a + i\Delta x$:

$$x_0 = a, x_1 = a + \Delta x, \ldots, x_{n-1} = a + (n-1)\Delta x, x_n = b.$$

Then the integral $\int_a^b f(x)\,dx$ can be approximated by

$$\int_a^b f(x)\,dx \approx \frac{\Delta x}{3}\left(f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \cdots + 4f(x_{n-1}) + f(x_n)\right).$$

As with the trapezoid method, this is useful only with an **error bound**:

Theorem 2.49: Error for Simpson's Approximation

Suppose f has a fourth derivative $f^{(4)}$ everywhere on the interval $[a,b]$, and $|f^{(4)}(x)| \le M$ for all x in the interval. With $\Delta x = (b-a)/n$, an error bound for Simpson's approximation is

$$E(\Delta x) = \frac{b-a}{180} M (\Delta x)^4 = \frac{(b-a)^5}{180 n^4} M.$$

Note: Because of the factor $180n^4$, the error bound is usually much more accurate than that for the Trapezoid Rule or the Midpoint Rule. And since the formula is only slightly more complicated, Simpson's Rule is extremely useful and probably one of the most commonly used integral approximation rules in practice.

Example 2.50: Approximate an Integral With Parabolas

Let us again approximate $\int_0^1 e^{-x^2}\,dx$ to two decimal places.

Solution. The fourth derivative of $f(x) = e^{-x^2}$ is $(16x^4 - 48x^2 + 12)e^{-x^2}$ and on $[0,1]$ this is at most 12 in absolute value by using a graphing calculator or computer software to estimate the maximum value. We begin by estimating the number of subintervals we are likely to need. To get two decimal places of accuracy, we will certainly need $E(\Delta x) < 0.005$, but taking a cue from our earlier example, let's require $E(\Delta x) < 0.001$:

$$\begin{aligned}
\frac{1}{180}(12)\frac{1}{n^4} &< 0.001 \\
\frac{200}{3} &< n^4 \\
2.86 \approx \sqrt[[4]]{\frac{200}{3}} &< n
\end{aligned}$$

So we try $n = 4$, since we need an even number of subintervals. Then the error bound is $12/180/4^4 < 0.0003$ and the approximation is

$$(f(0)+4f(1/4)+2f(1/2)+4f(3/4)+f(1))\frac{1}{3\cdot 4} \approx 0.746855.$$

So the true value of the integral is between $0.746855-0.0003 = 0.746555$ and $0.746855+0.0003 = 0.7471555$, both of which round to 0.75. This is a great approximation of the integral considering it only requires 4 evaluations of the integrand function! Remember that significantly more terms were required for the Trapezoid Rule approximation. ♣

Exercises for Section 2.6

Exercise 2.6.1 *Approximate the following integrals by computing the Trapezoid and Simpson approximations using 4 subintervals, and compute the error bound for each. (Finding the maximum values of the second and fourth derivatives can be challenging for some of these; you may use a graphing calculator or computer software to estimate the maximum values.)*

(a) $\displaystyle\int_1^3 x\,dx$

(b) $\displaystyle\int_0^3 x^2\,dx$

(c) $\displaystyle\int_2^4 x^3\,dx$

(d) $\displaystyle\int_1^3 \frac{1}{x}\,dx$

(e) $\displaystyle\int_1^2 \frac{1}{1+x^2}\,dx$

(f) $\displaystyle\int_0^1 x\sqrt{1+x}\,dx$

(g) $\displaystyle\int_1^5 \frac{x}{1+x}\,dx$

(h) $\displaystyle\int_0^1 \sqrt{x^3+1}\,dx$

(i) $\displaystyle\int_0^1 \sqrt{x^4+1}\,dx$

(j) $\displaystyle\int_1^4 \sqrt{1+1/x}\,dx$

Exercise 2.6.2 *Using Simpson's Rule on a parabola $f(x)$, even with just two subintervals, gives the exact value of the integral, because the parabolas used to approximate f will be f itself. Remarkably, Simpson's Rule also computes the integral of a cubic function $f(x) = ax^3+bx^2+cx+d$ exactly. Show this is true by showing that*

$$\int_{x_0}^{x_2} f(x)\,dx = \frac{x_2-x_0}{3\cdot 2}(f(x_0)+4f((x_0+x_2)/2)+f(x_2)).$$

This does require a bit of messy algebra, so you may prefer to use Sage.

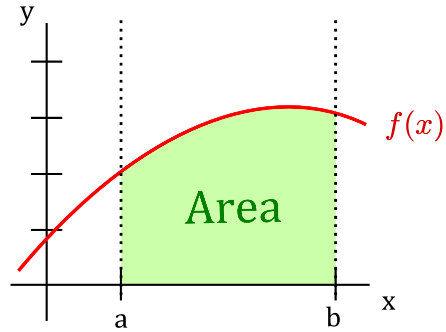
2.7 Improper Integrals

Recall that the Fundamental Theorem of Calculus says that if f is a **continuous** function on the **closed interval** $[a,b]$, then

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

where F is any antiderivative of f.

Both the **continuity** condition and **closed interval** must hold to use the Fundamental Theorem of Calculus, and in this case, $\int_a^b f(x)\,dx$ represents the net area under $f(x)$ from a to b:



We begin with an example where blindly applying the Fundamental Theorem of Calculus can give an incorrect result.

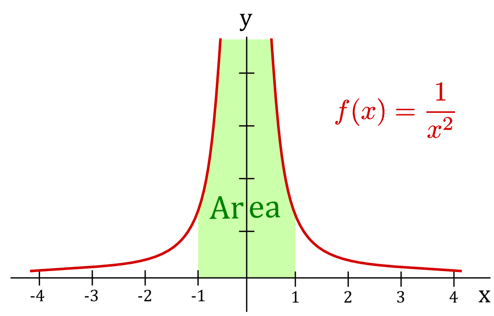
Example 2.51: Using FTC

Explain why $\int_{-1}^{1} \frac{1}{x^2}\,dx$ *is not equal to* -2.

Solution. Here is how one might proceed:

$$\int_{-1}^{1} \frac{1}{x^2}\,dx = \int_{-1}^{1} x^{-2}\,dx = -x^{-1}\Big|_{-1}^{1} = -\frac{1}{x}\Big|_{-1}^{1} = \left(-\frac{1}{1}\right)-\left(-\frac{1}{(-1)}\right) = -2$$

However, the above answer is **WRONG!** Since $f(x) = 1/x^2$ is not continuous on $[-1,\ 1]$, we cannot directly apply the Fundamental Theorem of Calculus. Intuitively, we can see why -2 is not the correct answer by looking at the graph of $f(x) = 1/x^2$ on $[-1,\ 1]$. The shaded area appears to grow without bound as seen in the figure below.



♣

Formalizing this example leads to the concept of an improper integral. There are two ways to extend the Fundamental Theorem of Calculus. One is to use an **infinite interval**, i.e., $[a,\infty)$, $(-\infty,b]$ or $(-\infty,\infty)$. The second is to allow the interval $[a,b]$ to contain an infinite **discontinuity** of $f(x)$. In either case, the integral is called an **improper integral**. One of the most important applications of this concept is probability distributions because determining quantities like the cumulative distribution or expected value typically require integrals on infinite intervals.

2.7.1. Improper Integrals: Infinite Limits of Integration

To compute improper integrals, we use the concept of limits along with the Fundamental Theorem of Calculus.

Definition 2.52: Improper Integrals – One Infinite Limit of Integration

If $f(x)$ is continuous on $[a,\infty)$, then the improper integral of f over $[a,\infty)$ is

$$\int_a^\infty f(x)\,dx = \lim_{R\to\infty}\int_a^R f(x)\,dx.$$

If $f(x)$ is continuous on $(-\infty,b]$, then the improper integral of f over $(-\infty,b]$ is

$$\int_{-\infty}^b f(x)\,dx = \lim_{R\to-\infty}\int_R^b f(x)\,dx.$$

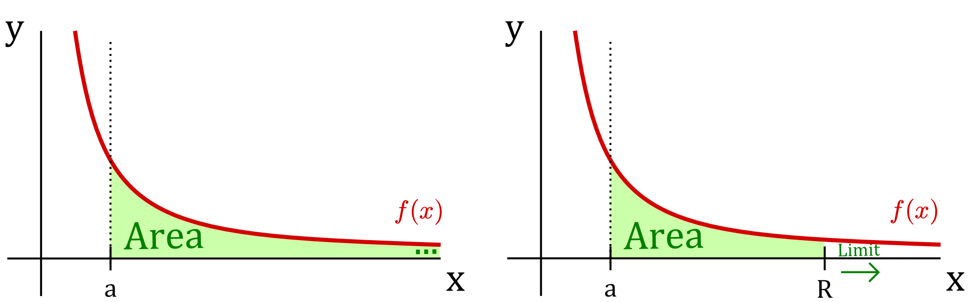
Since we are dealing with limits, we are interested in convergence and divergence of the improper integral. If the limit exists and is a finite number, we say the improper integral **converges**. Otherwise, we say the improper integral **diverges**, which we capture in the following definition.

Definition 2.53: Convergence and Divergence

If the limit exists and is a finite number, we say the improper integral **converges***.*

If the limit is $\pm\infty$ or does not exist, we say the improper integral **diverges***.*

To get an intuitive (though not completely correct) interpretation of improper integrals, we attempt to analyze $\int_a^\infty f(x)\,dx$ graphically. Here assume $f(x)$ is continuous on $[a,\infty)$:



We let R be a fixed number in $[a,\infty)$. Then by taking the limit as R approaches ∞, we get the improper integral:

$$\int_a^\infty f(x)\,dx = \lim_{R\to\infty}\int_a^R f(x)\,dx.$$

We can then apply the Fundamental Theorem of Calculus to the last integral as $f(x)$ is continuous on the closed interval $[a,R]$.

We next define the improper integral for the interval $(-\infty,\ \infty)$.

Definition 2.54: Improper Integrals – Two Infinite Limits of Integration

If both $\int_{-\infty}^{a} f(x)\,dx$ *and* $\int_{a}^{\infty} f(x)\,dx$ *are convergent, then the improper integral of* f *over* $(-\infty,\infty)$ *is*

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{a} f(x)\,dx + \int_{a}^{\infty} f(x)\,dx$$

The above definition requires **both** of the integrals

$$\int_{-\infty}^{a} f(x)\,dx \qquad \text{and} \qquad \int_{a}^{\infty} f(x)\,dx$$

to be convergent for $\int_{-\infty}^{\infty} f(x)\,dx$ to also be convergent. If **either** of $\int_{-\infty}^{a} f(x)\,dx$ or $\int_{a}^{\infty} f(x)\,dx$ is divergent, then so is $\int_{-\infty}^{\infty} f(x)\,dx$.

Example 2.55: Improper Integral – One Infinite Limit of Integration

Determine whether $\int_{1}^{\infty} \frac{1}{x}\,dx$ *is convergent or divergent.*

Solution. Using the definition for improper integrals we write this as:

$$\int_{1}^{\infty} \frac{1}{x}\,dx = \lim_{R\to\infty} \int_{1}^{R} \frac{1}{x}\,dx = \lim_{R\to\infty} \ln|x|\Big|_{1}^{R} = \lim_{R\to\infty} \ln|R| - \ln|1| = \lim_{R\to\infty} \ln|R| = +\infty$$

Therefore, the integral is **divergent**. ♣

Example 2.56: Improper Integral – Two Infinite Limits of Integration

Determine whether $\int_{-\infty}^{\infty} x\sin(x^2)\,dx$ *is convergent or divergent.*

Solution. We must compute both $\int_{0}^{\infty} x\sin(x^2)\,dx$ and $\int_{-\infty}^{0} x\sin(x^2)\,dx$. Note that we don't have to split the integral up at 0, any finite value a will work. First we compute the indefinite integral. Let $u = x^2$, then $du = 2x\,dx$ and hence,

$$\int x\sin(x^2)\,dx = \frac{1}{2}\int \sin(u)\,du = -\frac{1}{2}\cos(x^2) + C$$

Using the definition of improper integral gives:

$$\int_{0}^{\infty} x\sin(x^2)\,dx = \lim_{R\to\infty} \int_{0}^{R} x\sin(x^2)\,dx = \lim_{R\to\infty} \left[-\frac{1}{2}\cos(x^2)\right]\Bigg|_{0}^{R} = -\frac{1}{2}\lim_{R\to\infty} \cos(R^2) + \frac{1}{2}$$

This limit does not exist since $\cos x$ **oscillates** between -1 and $+1$. In particular, $\cos x$ does not approach any particular value as x gets larger and larger. Thus, $\int_{0}^{\infty} x\sin(x^2)\,dx$ diverges, and hence, $\int_{-\infty}^{\infty} x\sin(x^2)\,dx$ diverges. ♣

2.7.2. Improper Integrals: Discontinuities

When there is a discontinuity in $[a,b]$ or at an endpoint, then the improper integral is as follows.

Definition 2.57: Improper Integrals – Discontinuities on Integration Bounds

If $f(x)$ is continuous on $(a,b]$, then the improper integral of f over $(a,b]$ is

$$\int_a^b f(x)\,dx = \lim_{R\to a^+} \int_R^b f(x)\,dx.$$

If $f(x)$ is continuous on $[a,b)$, then the improper integral of f over $[a,b)$ is

$$\int_a^b f(x)\,dx = \lim_{R\to b^-} \int_a^R f(x)\,dx.$$

Definition 2.53 on convergence and divergence of an improper integral holds here as well: If the limit above exists and is a finite number, we say the improper integral **converges**. Otherwise, we say the improper integral **diverges**.

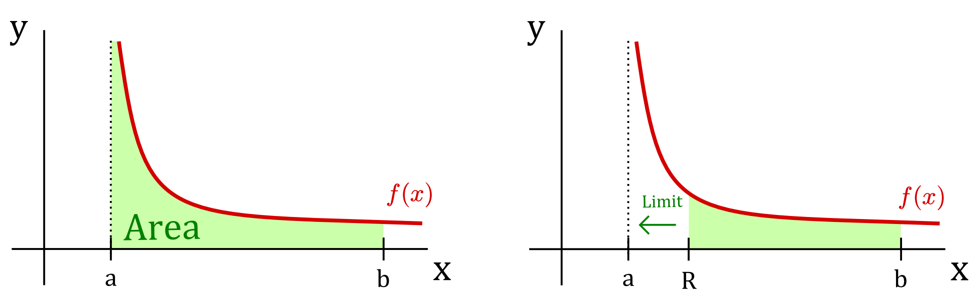
When there is a discontinuity in the interior of $[a,b]$, we use the following definition.

Definition 2.58: Improper Integrals – Discontinuities Within Integration Interval

If f has a discontinuity at $x=c$ where $c\in[a,b]$, and both $\int_a^c f(x)\,dx$ and $\int_c^b f(x)\,dx$ are convergent, then f over $[a,b]$ is

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

Again, we can get an intuitive sense of this concept by analyzing $\int_a^b f(x)\,dx$ graphically. Here assume $f(x)$ is continuous on $(a,b]$ but discontinuous at $x=a$:



We let R be a fixed number in (a,b). Then by taking the limit as R approaches a from the **right**, we get the improper integral:

$$\int_a^b f(x)\,dx = \lim_{R\to a^+} \int_R^b f(x)\,dx.$$

Now we can apply FTC to the last integral as $f(x)$ is continuous on $[R,b]$.

Example 2.59: A Divergent Integral

Determine if $\int_{-1}^{1} \frac{1}{x^2}\,dx$ *is convergent or divergent.*

Solution. The function $f(x) = 1/x^2$ has a discontinuity at $x = 0$, which lies in $[-1,1]$. We must compute $\int_{-1}^{0} \frac{1}{x^2}\,dx$ and $\int_{0}^{1} \frac{1}{x^2}\,dx$. Let's start with $\int_{0}^{1} \frac{1}{x^2}\,dx$:

$$\int_0^1 \frac{1}{x^2}\,dx = \lim_{R\to 0^+} \int_R^1 \frac{1}{x^2}\,dx = \lim_{R\to 0^+} -\frac{1}{x}\bigg|_R^1 = -1 + \lim_{R\to 0^+} \frac{1}{R}$$

which diverges to $+\infty$. Therefore, $\int_{-1}^{1} \frac{1}{x^2}\,dx$ is **divergent** since one of $\int_{-1}^{0} \frac{1}{x^2}\,dx$ and $\int_{0}^{1} \frac{1}{x^2}\,dx$ is divergent.

♣

Example 2.60: Integral of the Logarithm

Determine if $\int_0^1 \ln x\,dx$ *is convergent or divergent. Evaluate it if it is convergent.*

Solution. Note that $f(x) = \ln x$ is discontinuous at the endpoint $x = 0$. We first use Integration by Parts to compute $\int \ln x\,dx$. We let $u = \ln x$ and $dv = dx$. Then $du = (1/x)dx$, $v = x$, giving:

$$\begin{aligned}\int \ln x\,dx &= x\ln x - \int x \cdot \frac{1}{x}\,dx \\ &= x\ln x - \int 1\,dx \\ &= x\ln x - x + C\end{aligned}$$

Now using the definition of improper integral for $\int_0^1 \ln x\,dx$:

$$\int_0^1 \ln x\,dx = \lim_{R\to 0^+} \int_R^1 \ln x\,dx = \lim_{R\to 0^+} (x\ln x - x)\bigg|_R^1 = -1 - \lim_{R\to 0^+} (R\ln R) + \lim_{R\to 0^+} R$$

Note that $\lim_{R\to 0^+} R = 0$. We next compute $\lim_{R\to 0^+} (R\ln R)$. First, we rewrite the expression as follows:

$$\lim_{x\to 0^+} (R\ln R) = \lim_{R\to 0^+} \frac{\ln R}{1/R}.$$

Now the limit is of the indeterminate type $(-\infty)/(\infty)$ and l'Hôpital's Rule can be applied.

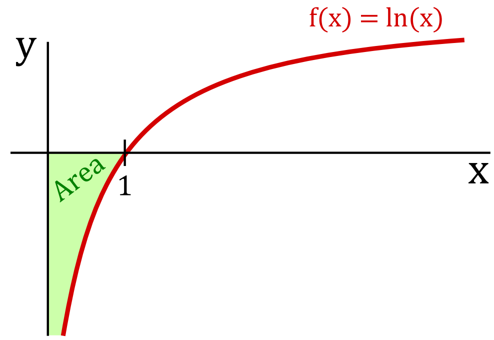
$$\lim_{R\to 0^+} (R\ln R) = \lim_{R\to 0^+} \frac{\ln R}{1/R} = \lim_{R\to 0^+} \frac{1/R}{-1/R^2} = \lim_{R\to 0^+} -\frac{R^2}{R} = \lim_{R\to 0^+} (-R) = 0$$

Thus, $\lim_{R\to 0^+}(R\ln R) = 0$. Thus

$$\int_0^1 \ln x\,dx = -1,$$

and the integral is convergent to -1.

Graphically, one might interpret this to mean that the net area under $\ln x$ on $[0,1]$ is -1 (the area in this case lies below the x-axis).



♣

Example 2.61: Integral of a Square Root

Determine if $\int_0^4 \frac{dx}{\sqrt{4-x}}$ *is convergent or divergent. Evaluate it if it is convergent.*

Solution. Note that $\frac{1}{\sqrt{4-x}}$ is discontinuous at the endpoint $x = 4$. We use a u-substitution to compute $\int \frac{dx}{\sqrt{4-x}}$. We let $u = 4-x$, then $du = -dx$, giving:

$$\begin{aligned}\int \frac{dx}{\sqrt{4-x}} &= \int -\frac{du}{u^{1/2}}\\ &= \int -u^{-1/2}\,du\\ &= -2(u)^{1/2} + C\\ &= -2\sqrt{4-x} + C\end{aligned}$$

Now using the definition of improper integrals for $\int_0^4 \frac{dx}{\sqrt{4-x}}$:

$$\int_0^4 \frac{dx}{\sqrt{4-x}} = \lim_{R\to 4^-}\left(-2\sqrt{4-x}\right)\Big|_0^R = \lim_{R\to 4^-} -2\sqrt{4-R} + 2\sqrt{4} = 4$$

♣

Example 2.62: Improper Integral

Determine if $\int_1^2 \frac{dx}{(x-1)^{1/3}}$ *is convergent or divergent. Evaluate it if it is convergent.*

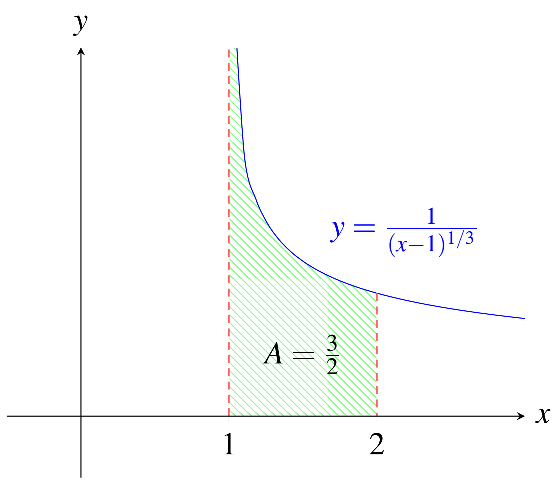
Solution. Note that $f(x) = \dfrac{1}{(x-1)^{1/3}}$ is discontinuous at the endpoint $x = 1$. We first use substitution to find $\int \dfrac{dx}{(x-1)^{1/3}}$. We let $u = x - 1$. Then $du = dx$, giving

$$\int \frac{dx}{(x-1)^{1/3}} = \int \frac{du}{u^{1/3}} = \int u^{-1/3}du = \frac{3}{2}u^{2/3} + C = \frac{3}{2}(x-1)^{2/3} + C.$$

Now using the definition of improper integral for $\int_1^2 \dfrac{dx}{(x-1)^{1/3}}$:

$$\int_1^2 \frac{dx}{(x-1)^{1/3}} = \lim_{R\to 1^+} \int_R^2 \frac{dx}{(x-1)^{1/3}} = \lim_{R\to 1^+} \frac{3}{2}(x-1)^{2/3}\bigg|_R^2 = \frac{3}{2} - \lim_{R\to 1^+} \frac{3}{2}(R-1)^{2/3} = \frac{3}{2},$$

and the integral is convergent to $\frac{3}{2}$. Graphically, one might interpret this to mean that the net area under $\dfrac{1}{(x-1)^{1/3}}$ on $[1,2]$ is $\frac{3}{2}$.



♣

2.7.3. p-Integrals

Integrals of the form $\dfrac{1}{x^p}$ come up again in the study of series. These integrals can be either classified as an improper integral with an infinite limit of integration, $\int_a^\infty \dfrac{1}{x^p}\,dx$, or as an improper integral with discontinuity at $x = 0$, $\int_0^a \dfrac{1}{x^p}\,dx$. In asymptotic analysis, it is useful to know when either of these intervals converge or diverge.

Theorem 2.63: p-Test for Infinite Limit

For $a > 0$:

(i) If $p > 1$, then $\int_a^\infty \frac{1}{x^p}\, dx$ **converges**.

(ii) If $p \leq 1$, then $\int_a^\infty \frac{1}{x^p}\, dx$ **diverges**.

Proof.

(i) If $p > 1$, we have $\int_a^\infty \frac{1}{x^p}\, dx = \lim_{R\to\infty} \frac{x^{1-p}}{1-p}\bigg|_a^R = \lim_{R\to\infty} \frac{R^{1-p}}{1-p} - \frac{a^{1-p}}{1-p} = \frac{a^{1-p}}{p-1}$.

(ii) If $p \leq 1$, the above tells us that the resulting limit is infinite.

♣

Theorem 2.64: p-Test for Discontinuity

For $a > 0$:

(i) If $p < 1$, then $\int_0^a \frac{1}{x^p}\, dx$ **converges**.

(ii) If $p \geq 1$, then $\int_0^a \frac{1}{x^p}\, dx$ **diverges**.

Proof.

(i) If $p < 1$, we have to $\int_0^a \frac{1}{x^p}\, dx = \lim_{R\to 0^+} \frac{x^{1-p}}{1-p}\bigg|_R^a = \lim_{R\to 0^+} \frac{a^{1-p}}{1-p} - \frac{R^{1-p}}{1-p} = \frac{a^{1-p}}{1-p}$.

(ii) If $p \geq 1$, the above tells us that the resulting limit is infinite.

♣

With Example 2.55 and Example 2.59, you have already seen how the p-Test is applied. For good measure, here is one more example.

Example 2.65: p-Test

Determine if the following integrals are convergent or divergent.

(a) $\int_1^\infty \frac{1}{x^3}\, dx$ *(b)* $\int_0^5 \frac{1}{x^4}\, dx$

Solution.

(a) This is a p-integral with an infinite upper limit of integration and $p = 3 > 1$. Therefore, by the p-Test for Infinite Limit, $\int_1^\infty \frac{1}{x^3}\,dx$ converges.

(b) We classify $\int_0^5 \frac{1}{x^4}\,dx$ as a p-integral with a discontinuity at $x = 0$ and $p = 4 \geq 1$. Thus, by the p-Test for Discontinuity, the integral diverges.

♣

2.7.4. Comparison Test

The following test allows us to determine convergence/divergence information about improper integrals that are hard to compute by comparing them to easier ones. We state the test for $[a, \infty)$, but similar versions hold for the other improper integrals.

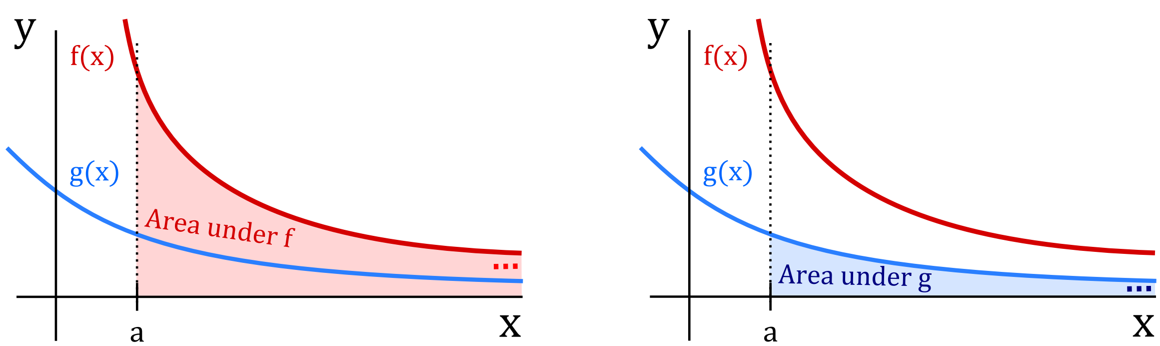
Theorem 2.66: Comparison Test for Improper Integrals

Assume that $f(x) \geq g(x) \geq 0$ *for* $x \geq a$.

(i) If $\int_a^\infty f(x)\,dx$ **converges**, *then* $\int_a^\infty g(x)\,dx$ *also* **converges**.

(ii) If $\int_a^\infty g(x)\,dx$ **diverges**, *then* $\int_a^\infty f(x)\,dx$ *also* **diverges**.

Informally, (i) says that if $f(x)$ is larger than $g(x)$, and the area under $f(x)$ is finite (converges), then the area under $g(x)$ must also be finite (converges). Informally, (ii) says that if $f(x)$ is larger than $g(x)$, and the area under $g(x)$ is infinite (diverges), then the area under $f(x)$ must also be infinite (diverges).



Example 2.67: Comparison Test

Show that $\int_2^\infty \frac{\cos^2 x}{x^2}\,dx$ *converges.*

Solution. We use the Comparison Test to show that it converges. Note that $0 \leq \cos^2 x \leq 1$ and hence

$$0 \leq \frac{\cos^2 x}{x^2} \leq \frac{1}{x^2}.$$

Thus, taking $f(x) = 1/x^2$ and $g(x) = \cos^2 x/x^2$ we have $f(x) \geq g(x) \geq 0$. One can easily see that $\int_2^\infty \frac{1}{x^2}\,dx$ converges. Therefore, $\int_2^\infty \frac{\cos^2 x}{x^2}\,dx$ also converges. ♣

Exercises for Section 2.7

Exercise 2.7.1 *Determine whether the following improper integrals are convergent or divergent. Evaluate those that are convergent.*

(a) $\displaystyle\int_0^\infty \frac{1}{x^2+1}\,dx$

(b) $\displaystyle\int_0^\infty \frac{x}{x^2+1}\,dx$

(c) $\displaystyle\int_0^\infty e^{-x}(\cos x + \sin x)\,dx.$

(d) $\displaystyle\int_0^{\pi/2} \sec^2 x\,dx$

(e) $\displaystyle\int_0^4 \frac{1}{(4-x)^{2/5}}\,dx$

(f) $\displaystyle\int_1^\infty \frac{1}{x^2}\,dx$

(g) $\displaystyle\int_e^\infty \frac{1}{x\sqrt{\ln x}}\,dx$

(h) $\displaystyle\int_0^\infty e^{-3x}\,dx$

(i) $\displaystyle\int_1^e \frac{1}{x(\ln x)^2}\,dx$

(j) $\displaystyle\int_0^\infty e^{-x}\sin^2\left(\frac{\pi x}{2}\right)\,dx$

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

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

Exercise 2.7.2 *Prove that the integral* $\int_1^\infty \frac{1}{x^p}\,dx$ *is convergent if* $p > 1$ *and divergent if* $0 < p \leq 1$.

Exercise 2.7.3 *Suppose that* $p > 0$. *Find all values of* p *for which* $\int_0^1 \frac{1}{x^p}\,dx$ *converges.*

Exercise 2.7.4 *Show that* $\int_1^\infty \frac{\sin^2 x}{x(\sqrt{x}+1)}\,dx$ *converges.*

2.8 Additional exercises

These problems require the techniques of this chapter, and are in no particular order. Some problems may be done in more than one way.

Exercise 2.8.1 $\displaystyle\int (t+4)^3\,dt$

Exercise 2.8.2 $\displaystyle\int t(t^2-9)^{3/2}\,dt$

Exercise 2.8.3 $\displaystyle\int (e^{t^2}+16)te^{t^2}\,dt$

Exercise 2.8.4 $\displaystyle\int \sin t\cos 2t\,dt$

Exercise 2.8.5 $\displaystyle\int \tan t\sec^2 t\,dt$

Exercise 2.8.6 $\displaystyle\int \frac{2t+1}{t^2+t+3}\,dt$

Exercise 2.8.7 $\displaystyle\int \frac{1}{t(t^2-4)}\,dt$

Exercise 2.8.8 $\displaystyle\int \frac{1}{(25-t^2)^{3/2}}\,dt$

Exercise 2.8.9 $\displaystyle\int \frac{\cos 3t}{\sqrt{\sin 3t}}\,dt$

Exercise 2.8.10 $\displaystyle\int t\sec^2 t\,dt$

Exercise 2.8.11 $\displaystyle\int \frac{e^t}{\sqrt{e^t+1}}\,dt$

Exercise 2.8.12 $\displaystyle\int \cos^4 t\,dt$

Exercise 2.8.13 $\displaystyle\int \frac{1}{t^2+3t}\,dt$

Exercise 2.8.14 $\displaystyle\int \frac{1}{t^2\sqrt{1+t^2}}\,dt$

Exercise 2.8.15 $\displaystyle\int \frac{\sec^2 t}{(1+\tan t)^3}\,dt$

Exercise 2.8.16 $\int t^3\sqrt{t^2+1}\,dt$

Exercise 2.8.17 $\int e^t \sin t\,dt$

Exercise 2.8.18 $\int (t^{3/2}+47)^3\sqrt{t}\,dt$

Exercise 2.8.19 $\int \frac{t^3}{(2-t^2)^{5/2}}\,dt$

Exercise 2.8.20 $\int \frac{1}{t(9+4t^2)}\,dt$

Exercise 2.8.21 $\int \frac{\arctan 2t}{1+4t^2}\,dt$

Exercise 2.8.22 $\int \frac{t}{t^2+2t-3}\,dt$

Exercise 2.8.23 $\int \sin^3 t\cos^4 t\,dt$

Exercise 2.8.24 $\int \frac{1}{t^2-6t+9}\,dt$

Exercise 2.8.25 $\int \frac{1}{t(\ln t)^2}\,dt$

Exercise 2.8.26 $\int t(\ln t)^2\,dt$

Exercise 2.8.27 $\int t^3 e^t\,dt$

Exercise 2.8.28 $\int \frac{t+1}{t^2+t-1}\,dt$

3. Applications of Integration

Being able to calculate the length of a curve, the area under a curve or between curves, surface areas of 3D objects, or volumes of 3D objects as shown in Figure 3.1 are important concepts in mathematics and will appear in many applications. In this chapter, we will concern ourselves with area and volume computations that use techniques from integral calculus. Length and surface area computations use similar techniques but will not be discussed here.

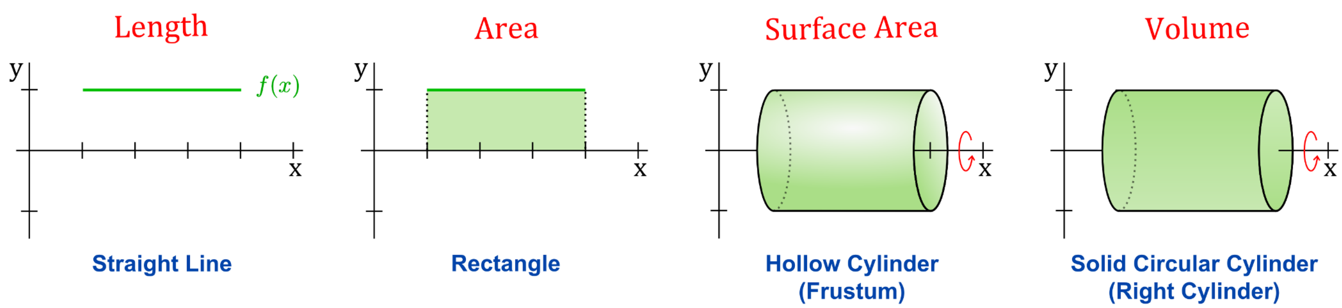


Figure 3.1

3.1 Average Value and Area Revisited

In this section we apply the tool of integration to two applications, namely that of finding the average value of a function and that of determining the area of a region bounded by functions.

3.1.1. Average Value of a Function

The average of some finite set of values is a familiar concept. If, for example, the class scores on a quiz are 10, 9, 10, 8, 7, 5, 7, 6, 3, 2, 7, 8, then the average score is the sum of these numbers divided by the size of the class:

$$\text{average score} = \frac{10+9+10+8+7+5+7+6+3+2+7+8}{12} = \frac{82}{12} \approx 6.83.$$

Although the above was a discrete example, we can extend the idea to that of continuous functions. If we are given an integrable function f on a closed interval $[a,b]$, then the average value of f on $[a,b]$ can be defined in a similar fashion. We begin by dividing the interval $[a,b]$ into n subintervals with equal width

$$\Delta x = \frac{b-a}{n}$$

and partition

$$P = \{x_1, x_2, \ldots, x_n\}$$

such that x_i is in the i-th subinterval. Then the average value f_{avg} over the n subintervals and choices x_i is given by

$$f_{avg} = \frac{f(x_1) + f(x_2) + \cdots + f(x_n)}{n} = \frac{\sum_{i=1}^{n} f(x_i)}{n}.$$

We now use a typical trick in mathematics, namely we multiply both sides by 1 in the form of $\frac{b-a}{b-a}$ and rearrange terms:

$$f_{avg} = \frac{b-a}{b-a} \cdot \frac{\sum_{x=1}^{n} f(x_i)}{n} = \frac{b-a}{n} \cdot \frac{\sum_{i=1}^{n} f(x_i)}{b-a}$$

Next, we replace $\frac{b-a}{n}$ by Δx to obtain

$$f_{avg} = \Delta x \cdot \frac{\sum_{i=1}^{n} f(x_i)}{b-a} = \frac{1}{b-a} \cdot \sum_{i=1}^{n} f(x_i)\Delta x.$$

Taking the limit as $n \to \infty$, we finally have

$$\begin{aligned} f_{avg} &= \lim_{n\to\infty} \frac{1}{b-a} \cdot \sum_{i=1}^{n} f(x_i)\Delta x \\ &= \frac{1}{b-a} \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x \\ &= \frac{1}{b-a} \int_a^b f(x)\,dx \end{aligned}$$

provided the limit exists.

We summarize this result in the following definition.

Definition 3.1: Average Value of a Function

If f is an integrable function on the closed interval $[a,b]$, then the **average value** f_{avg} *of f on $[a,b]$ is*

$$f_{avg} = \frac{1}{b-a} \int_a^b f(x)\,dx$$

provided the definite integral exists.

Example 3.2: Average Value Visualized

Suppose a function is defined by $f(x) = 16x^2 + 5$.

(a) What is the average value of f between $x = 1$ and $x = 3$?

(b) Interpret your result from part (a) geometrically.

Solution.

(a) By definition of the average value of a function, we have

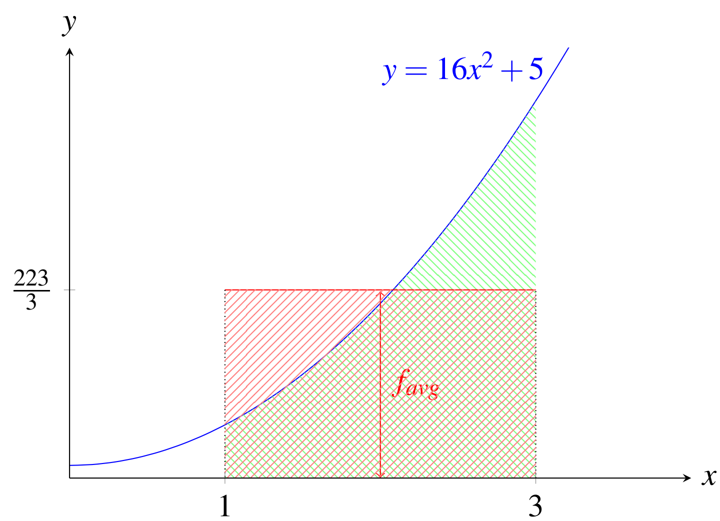
$$\begin{aligned} f_{avg} &= \frac{1}{3-1}\int_1^3 \left(16x^2+5\right)\,dx \\ &= \frac{1}{2}\left[\frac{16}{3x^3}+5x\right]_1^3 = \frac{1}{2}\left[\left(16\cdot 3^2+15\right)-\left(\frac{16\cdot 1^3}{3}+5\right)\right] \\ &= \frac{1}{2}\left(\frac{446}{3}\right) = \frac{223}{3} \end{aligned}$$

(b) We can interpret the same problem geometrically by asking the question: What is the average **height** of $f(x) = 16x^2+5$ on the interval $[1,3]$?

The area A under $f(x) = 16x^2+5$ on the interval $[1,3]$ is given by

$$A = \int_1^3 \left(16x^2+5\right)\,dx = \frac{446}{3}.$$

The area under the graph of $y = f_{avg} = \frac{223}{3}$ over the same interval $[1,3]$ is simply the area of a rectangle that is 2 units wide and 223/3 units high with area 446/3 units squared. So the average height of a function is is the height of the horizontal line that produces the same area over the given interval as shown below.

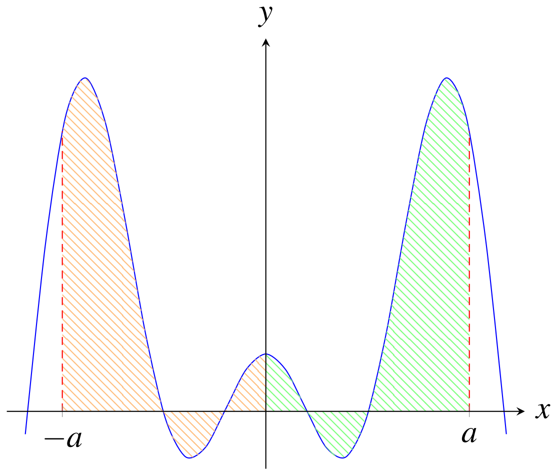


♣

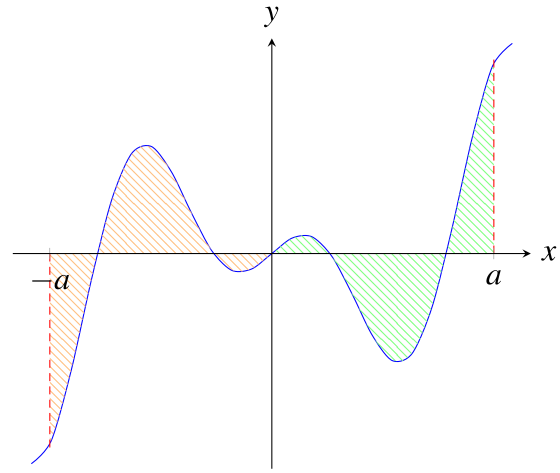
3.1.2. Area of Symmetric Functions

While we do not often work with even and odd functions, it is nonetheless useful to know the following facts about the area of the region under the curves of these functions on a symmetric interval $[-a,a]$. Figure 3.2a illustrates the area of the region under the curve of an even function f on the interval $[-a,a]$, and we are readily convinced that the net area shown is twice that of the net area of f on the interval $[0,a]$.

Similarly, Figure 3.2b illustrates the area of the region under the curve of an odd function f on the interval $[-a,a]$, and we readily believe that the net area is zero.



(a) f *even function*



(b) f *odd function*

Figure 3.2

Theorem 3.3: Integrals of Odd/Even Functions on $[-a,a]$

Let f be a continuous function on the symmetric interval $[-a,a]$.

1. *If f is even, then*
$$\int_{-a}^{a} f(x)\,dx = 2\int_{0}^{a} f(x)\,dx.$$
2. *If f is odd, then*
$$\int_{-a}^{a} f(x)\,dx = 0.$$

Proof. The proofs are readily accessible, so we only show the proof for item 1 of the theorem as the proof for item 2 is similar.

$$\begin{aligned}
\int_{-a}^{a} f(x)\,dx &= \int_{-a}^{0} f(x)\,dx + \int_{0}^{a} f(x)\,dx && \text{sum rule}\\
&= -\int_{0}^{-a} f(x)\,dx + \int_{0}^{a} f(x)\,dx && \text{order of limits rule}\\
&= \int_{0}^{a} f(-u)\,du + \int_{0}^{a} f(x)\,dx && u=-x,\ du=-dx,\ u(0)=0,\ u(-a)=a\\
&= \int_{0}^{a} f(u)\,du + \int_{0}^{a} f(x)\,dx && f \text{ even}\\
&= 2\int_{0}^{a} f(x)\,dx && \text{change dummy variables in first integral}
\end{aligned}$$

♣

3.1.3. Area Between Curves: Integration w.r.t x

We have seen how integration can be used to find an area between a curve and the x-axis. With very little change we can find some areas between curves; indeed, the area between a curve and the x-axis may be interpreted as the area between the curve and a second "curve" with equation $y = 0$.

Suppose we would like to find the area below $f(x) = -x^2 + 4x + 3$ and above $g(x) = -x^3 + 7x^2 - 10x + 5$ over the interval $1 \le x \le 2$. Figure 3.3a depicts the area of the region contained between the curves of f and g as well as the lines $x = 1$ and $x = 2$. We can see from this figure that the curve f is indeed above the curve g on the interval $[1, 2]$.

We can approximate the area between two curves by dividing the area into thin sections and approximating the area of each section by a rectangle, as indicated in Figure 3.3b. The area of a typical rectangle is $\Delta x(f(x_i) - g(x_i))$, so the total area is approximately

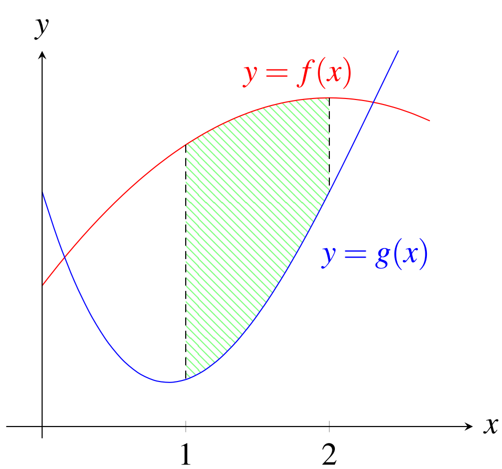
$$\sum_{i=0}^{n-1} (f(x_i) - g(x_i))\Delta x.$$

This is exactly the sort of sum that turns into an integral in the limit, namely the integral

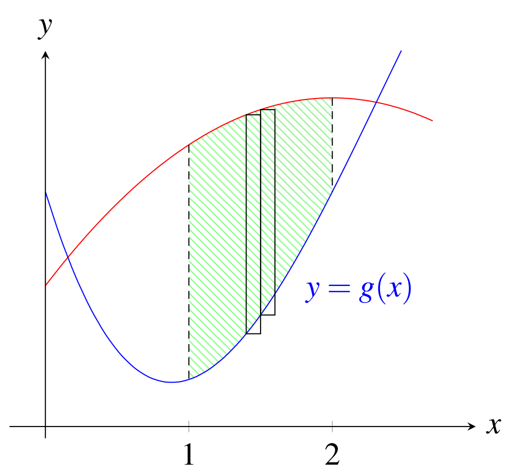
$$\int_1^2 f(x) - g(x)\, dx.$$

Then

$$\begin{aligned}
\int_1^2 f(x) - g(x)\, dx &= \int_1^2 (-x^2 + 4x + 3) - (-x^3 + 7x^2 - 10x + 5)\, dx \\
&= \int_1^2 x^3 - 8x^2 + 14x - 2\, dx = \left[\frac{x^4}{4} - \frac{8x^3}{3} + 7x^2 - 2x\right]_1^2 \\
&= \left(\frac{16}{4} - \frac{64}{3} + 28 - 4\right) - \left(\frac{1}{4} - \frac{8}{3} + 7 - 2\right) = 23 - \frac{56}{3} - \frac{1}{4} = \frac{49}{12}.
\end{aligned}$$



(a) *Area of region between f, g, $x = 1$ and $x = 2$*



(b) *Approximating area of region with rectangles*

Figure 3.3

We can also approach this problem of finding the area of the region between the curves of f and g as well as the lines $x=1$ and $x=2$ by using our knowledge of the definite integrals of f and g separately. Figure 3.4a depicts the area of the region between the curves f and the x-axis as well as the lines $x=1$ and $x=2$, which we know as the definite integral

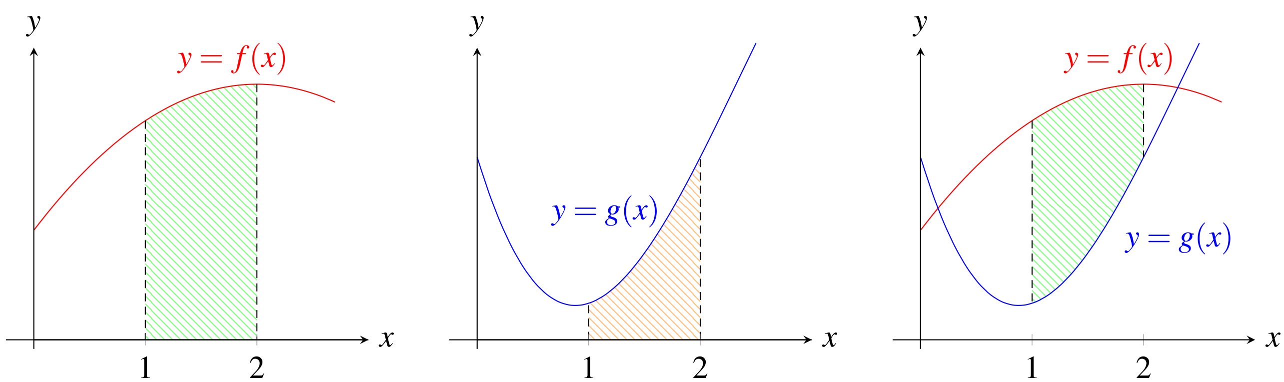
$$\int_1^2 f(x)\,dx.$$

Similarly, Figure 3.4b depicts the area of the region between the curves g and the x-axis as well as the lines $x=1$ and $x=2$, which we know as the definite integral

$$\int_1^2 g(x)\,dx.$$

From Figure 3.4, we can readily determine that the area of the region we want must be the area of the region under the graph of f minus the area of the region under the graph of g as shown again in Figure 3.4c. Once again, we have that

$$\int_1^2 f(x)-g(x)\,dx = \int_1^2 \left(-x^2+4x+3\right)-\left(-x^3+7x^2-10x+5\right)\,dx = \frac{49}{12}.$$



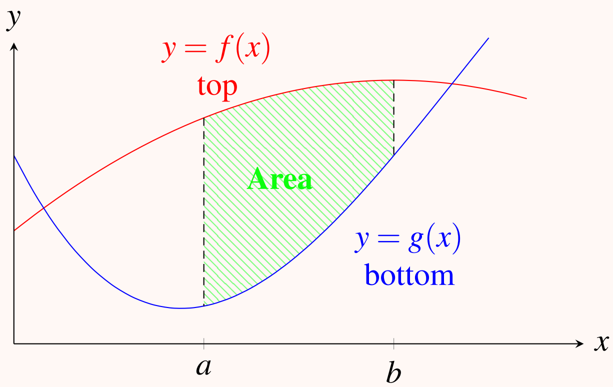
(a) $\int_1^2 f(x)\,dx$ (b) $\int_1^2 g(x)\,dx$ (c) $\int_1^2 [f(x)-g(x)]\,dx$

Figure 3.4

In summary, this procedure can informally be thought of as follows.

Area Between Two Curves: Integration w.r.t x

$$Area = \int_a^b (\text{top curve}) - (\text{bottom curve})\, dx, \qquad a \leq x \leq b.$$



This leads us to the following theorem for the area of the region between two curves on a closed interval.

Theorem 3.4: Area Between Two Curves: Integration w.r.t x

Let $y = f(x)$ and $y = g(x)$ be continuous functions on a closed interval $[a,b]$ such that $f(x) \geq g(x)$ for all $x \in [a,b]$. Then the area A of the region between the curves of f and g as well as the lines $x = a$ and $x = b$ is

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

Note: If we do not know which of the two functions $y = f(x)$ and $y = g(x)$ is larger on a given interval $[a,b]$, we can still ensure that we calculate the area of the region between the curves of f and g and the lines $x = a$ and $x = b$ by placing absolute values around the integrand:

$$A = \int_a^b |f(x) - g(x)|\, dx.$$

Example 3.5: Area Between Curves

Find the area between $f(x) = -x^2 + 4x$ and $g(x) = x^2 - 6x + 5$.

Solution. Here we are not given a specific interval, so it must be the case that there is a "natural" region involved, which we see from the graph shown below. Since the curves are both parabolas, the only reasonable interpretation is the region between the two intersection points.

We now find the intersection points by letting $f(x) = g(x)$:

$$\begin{aligned} -x^2 + 4x &= x^2 - 6x + 5 \\ -2x^2 + 10x - 5 &= 0 \\ \implies x &= \frac{5}{2} \pm \frac{\sqrt{15}}{2}. \end{aligned}$$

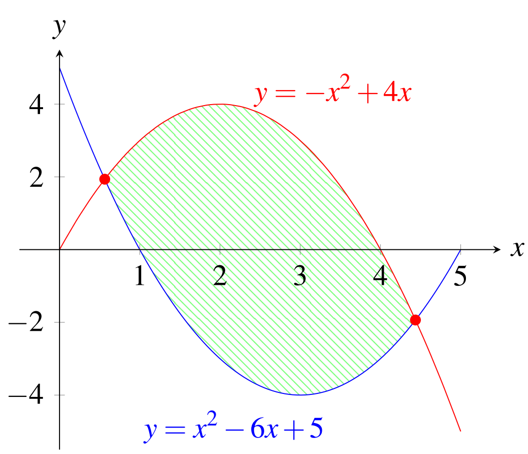
If we let

$$a = \frac{5 - \sqrt{15}}{2} \quad \text{and} \quad b = \frac{5 + \sqrt{15}}{2},$$

the total area is

$$\begin{aligned} \int_a^b \left(-x^2 + 4x\right) - (x^2 - 6x + 5)\,dx &= \int_a^b -2x^2 + 10x - 5\,dx \\ &= -\frac{2x^3}{3} + 5x^2 - 5x\Big|_a^b \\ &= 5\sqrt{15}. \end{aligned}$$

after a bit of simplification.



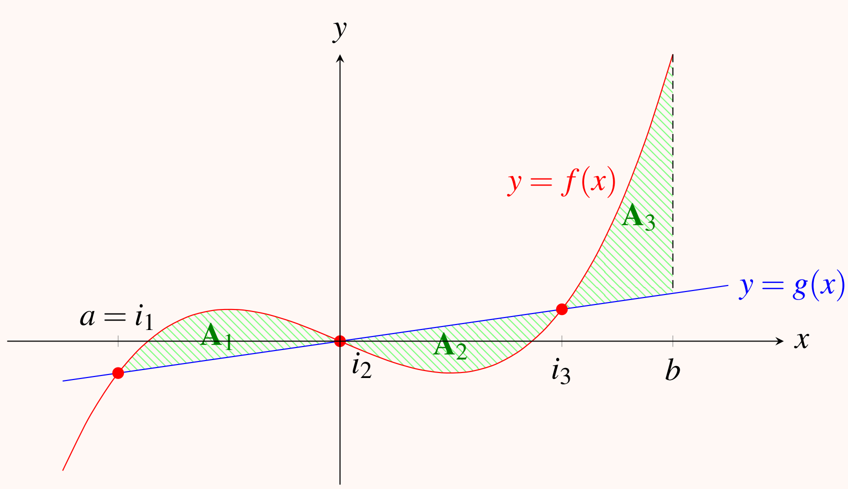
♣

A general guideline to compute the area between two curves follows.

Guideline for Finding Area Between Two Curves: Integration w.r.t x

Given two continuous functions $y = f(x)$ and $y = g(x)$ on a closed interval $[a,b]$, we can use the following steps to calculate the area of the region between the curves of f and g and the lines $x = a$ and $x = b$.

1. Find all intersection points of f and g in $[a,b]$. Suppose there are n of them called
$$i_1, i_2, \ldots, i_n.$$
2. Use these intersection points to create subintervals of $[a,b]$:
$$[a,b] = [a,i_1] \cup [i_1,i_2] \cup \cdots \cup [i_n,b].$$
3. Draw a sketch of the two curves on $[a,b]$ and indicate all intersection points, similar to the figure shown below.
4. Based on the sketch, determine top curve and the bottom curve in each subinterval.
5. For each subinterval, write the area of the region like so
$$\int_{\text{left bound}}^{\text{right bound}} [(\text{top curve}) - (\text{bottom curve})]\, dx$$
6. The area of the region between the curves f and g and the lines $x = a$ and $x = b$ is the sum of all intervals from step 5, similar to the area calculated below the figure.



$$\begin{aligned}\int_a^b |f(x) - g(x)|\, dx &= A_1 + A_2 + A_3 \\ &= \int_a^{i_2} (f(x) - g(x))\, dx + \int_{i_2}^{i_3} (g(x) - f(x))\, dx + \int_{i_3}^{b} (f(x) - g(x))\, dx\end{aligned}$$

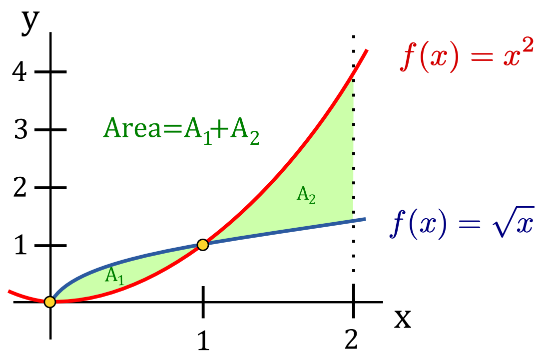
Example 3.6: Area Between Two Curves

Determine the area enclosed by $y = x^2$, $y = \sqrt{x}$, $x = 0$ *and* $x = 2$.

Solution. The points of intersection of $y = x^2$ and $y = \sqrt{x}$ are

$$x^2 = \sqrt{x} \implies x^4 = x \implies x^4 - x = 0 \implies x(x^3 - 1) = 0.$$

Thus, either $x = 0$ or $x = 1$. Sketching the curves gives:



The area we want to compute is the shaded region. Since the top curve changes at $x = 1$, we need to use the formula twice. For A_1 we have $a = 0$, $b = 1$, the top curve is $y = \sqrt{x}$ and the bottom curve is $y = x^2$. For A_2 we have $a = 1$, $b = 2$, the top curve is $y = x^2$ and the bottom curve is $y = \sqrt{x}$.

$$\textbf{Area} = \textbf{A}_1 + \textbf{A}_2 = \int_0^1 (\sqrt{x} - x^2)\,dx + \int_1^2 (x^2 - \sqrt{x})\,dx$$

For the first integral we have:

$$\int_0^1 (\sqrt{x} - x^2)\,dx = \left(\frac{2}{3}x^{3/2} - \frac{1}{3}x^3\right)\Big|_0^1 = \frac{1}{3}$$

Therefore, the second integral evaluates to

$$\int_1^2 (x^2 - \sqrt{x})\,dx = \left(\frac{1}{3}x^3 - \frac{2}{3}x^{3/2}\right)\Big|_1^2 = 3 - \frac{4\sqrt{2}}{3}$$

Thus,

$$\int_0^2 |x^2 - \sqrt{x}|\,dx = A_1 + A_2 = \frac{1}{3} + 3 - \frac{4\sqrt{2}}{3} = \frac{10 - 4\sqrt{2}}{3}.$$

♣

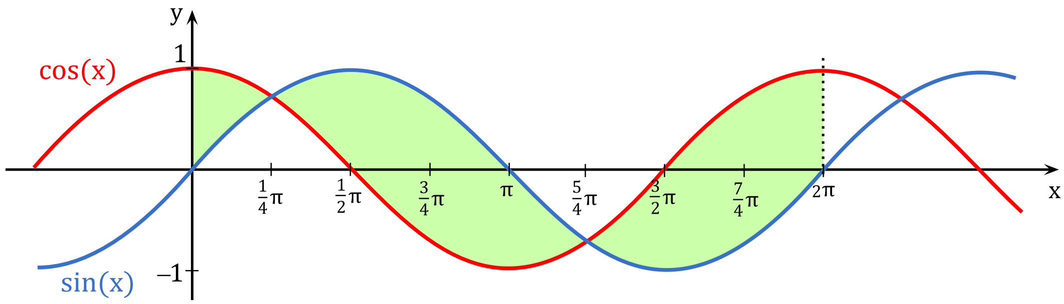
Example 3.7: Area Between Sine and Cosine

Determine the area enclosed by $y = \sin x$ *and* $y = \cos x$ *on the interval* $[0,\ 2\pi]$.

Solution. The curves $y = \sin x$ and $y = \cos x$ intersect when:

$$\sin x = \cos x \implies \tan x = 1 \implies x = \frac{\pi}{4} + \pi k,\ k \text{ an integer.}$$

We have the following sketch:



The area we want to compute is the shaded region. The top curve changes at $x = \pi/4$ and $x = 5\pi/4$, thus, we need to split the area up into three regions: from 0 to $\pi/4$; from $\pi/4$ to $5\pi/4$; and from $5\pi/4$ to 2π.

$$\begin{aligned}
\textbf{Area} &= \int_0^{\frac{\pi}{4}} (\cos x - \sin x)\,dx + \int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} (\sin x - \cos x)\,dx + \int_{\frac{5\pi}{4}}^{2\pi} (\cos x - \sin x)\,dx \\
&= (\sin x + \cos x)\Big|_0^{\pi/4} + (-\cos x - \sin x)\Big|_{\pi/4}^{5\pi/4} + (\sin x + \cos x)\Big|_{5\pi/4}^{2\pi} \\
&= \left(\sqrt{2} - 1\right) + \left(\sqrt{2} + \sqrt{2}\right) + \left(1 + \sqrt{2}\right) \\
&= 4\sqrt{2}
\end{aligned}$$

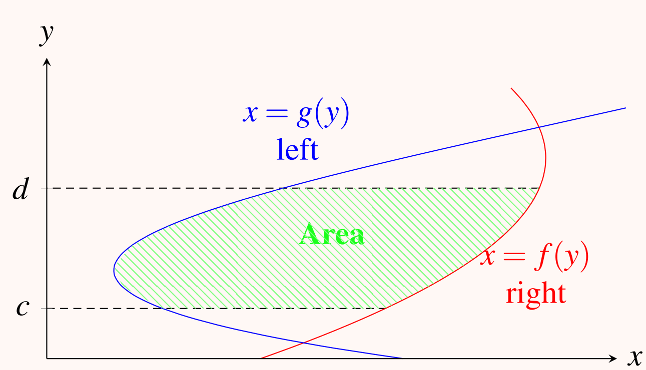
♣

3.1.4. Area Between Curves: Integration w.r.t. y

Sometimes the given curves are not functions of x. Instead, the area of the region is between two functions of y, namely $x = f(y)$ and $x = g(y)$, and the lines $y = c$ and $y = d$, which can be thought of informally as the following.

Area Between Two Curves: Integration w.r.t. y

$$Area = \int_c^d (\text{right curve}) - (\text{left curve})\,dy, \qquad c \le y \le d.$$



This leads us to a similar theorem as for the case when there are functions of x bounding a region.

Theorem 3.8: Area Between Two Curves: Integration w.r.t. y

Let $x = f(y)$ and $x = g(y)$ be two continuous functions on a closed interval $[c,d]$ such that $f(y) \geq g(y)$ for all y in $[c,d]$. Then the area A of the region between the curves of f and g as well as the lines $y = c$ and $y = d$ is

$$A = \int_c^d (f(y) - g(y))\, dy.$$

Note:

1. As before, if we do not know which of two functions $x = f(y)$ and $x = g(y)$ is larger on a given interval $[c,d]$, we can still ensure that we calculate the area of the region between the curves of f and g and the lines $y = c$ and $y = d$ by placing absolute values around the integrand:

$$\int_c^d |f(y) - g(y)|\, dy$$

2. Likewise, the *Guideline for Finding Area Between Two Curves: Integration w.r.t. x* can be adopted for this scenario as well, which we leave to the reader to do.

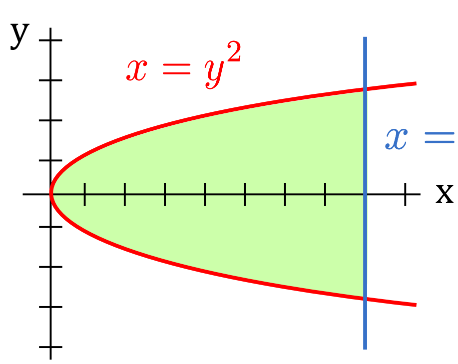
Example 3.9: Area Between Two Curves

Determine the area enclosed by $x = y^2$ and $x = 8$.

Solution. Note that $x = y^2$ and $x = 8$ intersect when:

$$y^2 = 8 \implies y = \pm\sqrt{8} \implies y = \pm 2\sqrt{2}$$

Sketching the two curves gives:



From the sketch $c = -2\sqrt{2}$, $d = 2\sqrt{2}$, the right curve is $x = 8$ and the left curve is $x = y^2$.

$$\textbf{Area} = \int_c^d [\text{right} - \text{left}]\, dy = \int_{-2\sqrt{2}}^{2\sqrt{2}} (8 - y^2)\, dy = \left(8y - \frac{1}{3}y^3\right)\Bigg|_{-2\sqrt{2}}^{2\sqrt{2}}$$

$$= \left[8(2\sqrt{2}) - \frac{1}{3}(2\sqrt{2})^3\right] - \left[8(-2\sqrt{2}) - \frac{1}{3}(-2\sqrt{2})^3\right] = \frac{64\sqrt{2}}{3}$$

We could have also integrated w.r.t. x by using the two functions $y = \sqrt{x}$ and $y = -\sqrt{x}$. We leave it to the reader to follow up on this integration. ♣

Exercises for Section 3.1

Exercise 3.1.1 *Find the average height of $f(x)$ over the given interval(s).*

(a) $f(x) = \cos x$ *on* $[0, \pi/2]$, $[-\pi/2, \pi/2]$, *and* $[0, 2\pi]$

(b) $f(x) = x^2$ *on* $[-2, 2]$

(c) $f(x) = \dfrac{1}{x^2}$ *on* $[1, a]$

(d) $f(x) = \sqrt{1 - x^2}$ *on* $[-1, 1]$

Exercise 3.1.2 *Find the area bounded by the curves by integrating with respect to x.*

(a) $y = x^4 - x^2$ *and* $y = x^2$ *(the part to the right of the y-axis)*

(b) $y = \cos(\pi x/2)$ *and* $y = 1 - x^2$ *(in the first quadrant)*

(c) $y = \sin(\pi x/3)$ *and* $y = x$ *(in the first quadrant)*

(d) $y = \sqrt{x}$ *and* $y = x^2$

(e) $y = x^{3/2}$ *and* $y = x^{2/3}$

(f) $y = x^2 - 2x$ *and* $y = x - 2$

Exercise 3.1.3 *Find the area bounded by the curves by integrating with respect to y.*

(a) $x = y^3$ *and* $x = y^2$

(b) $x = 0$ *and* $x = 25 - y^2$

Exercise 3.1.4 *Find the area bounded by the curves.*

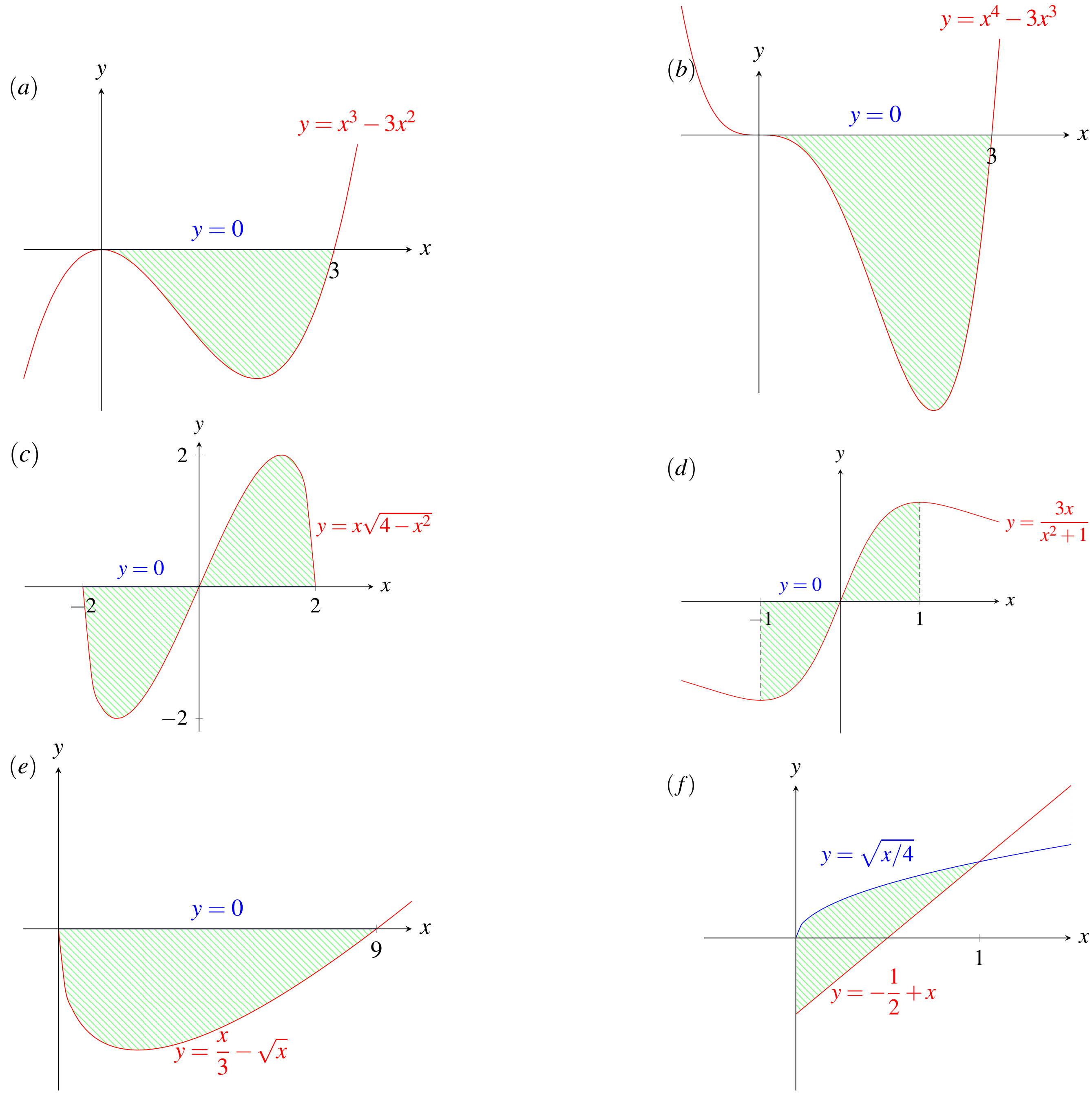
(a) $x = 1 - y^2$ *and* $y = -x - 1$

(b) $x = 3y - y^2$ *and* $x + y = 3$

(c) $y = \sqrt{x}$ *and* $y = \sqrt{x + 1}$, $0 \le x \le 4$

(d) $y = \sin x \cos x$ *and* $y = \sin x$, $0 \le x \le \pi$

Exercise 3.1.5 *Find the area bounded by the given curves.*



Exercise 3.1.6 *An object moves so that its velocity at time* t *is* $v(t) = -9.8t + 20$ *m/s. Describe the motion of the object between* $t = 0$ *and* $t = 5$*, find the total distance travelled by the object during that time, and find the net distance travelled.*

Exercise 3.1.7 *An object moves so that its velocity at time* t *is* $v(t) = \sin t$*. Set up and evaluate a single definite integral to compute the net distance travelled between* $t = 0$ *and* $t = 2\pi$*.*

Exercise 3.1.8 *An object moves so that its velocity at time t is $v(t) = 1 + 2\sin t$ m/s. Find the net distance travelled by the object between $t = 0$ and $t = 2\pi$, and find the total distance travelled during the same period.*

Exercise 3.1.9 *Consider the function $f(x) = (x+2)(x+1)(x-1)(x-2)$ on $[-2,2]$. Find the total area between the curve and the x-axis (measuring all area as positive).*

Exercise 3.1.10 *Consider the function $f(x) = x^2 - 3x + 2$ on $[0,4]$. Find the total area between the curve and the x-axis (measuring all area as positive).*

Exercise 3.1.11 *Evaluate the three integrals:*

$$A = \int_0^3 (-x^2+9)\,dx \qquad B = \int_0^4 (-x^2+9)\,dx \qquad C = \int_4^3 (-x^2+9)\,dx,$$

and verify that $A = B + C$. Draw a diagram which illustrates this relationship.

3.2 Applications to Business and Economics

Just like the process of differentiation is a useful tool in many business and economics applications such as problems related to elasticity of demand and optimization, so is the process of *anti*differentiation or integration. In this section, we revisit two important business and economic models, namely concerning law of supply and demand in a free-market environment and continuous money flow, and introduce integration as a method for solving problems of this nature.

3.2.1. Surplus in Consumption and Production

Recall that in a free market, the consumer demand for a particular commodity is dependent on the commodity's unit price, which is captured graphically by the demand curve. As expected, the quantity demanded of a commodity increases as the commodity's unit price decreases, and vice versa. Similarly, the unit price of a commodity is dependent on the commodity's availability in the market, which is articulated graphically by the supply curve. Typically, an increase or decrease in the commodity's unit price induces the producer to respectively increase or decrease the supply of the commodity. Below we see the supply and demand curves of a certain item produced and sold:

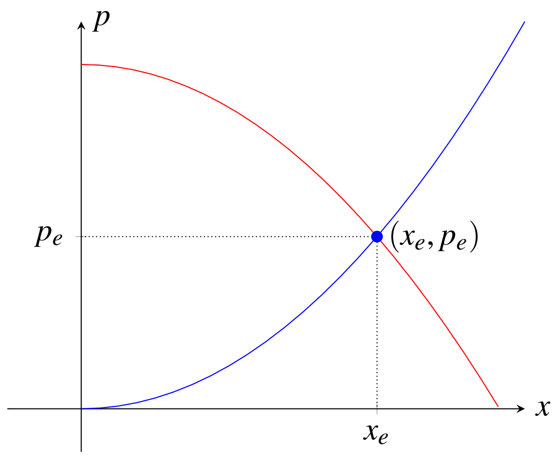


Figure 3.5: Example of a supply curve (in blue) and a demand curve (in red). The point of intersection (x_e, p_e) corresponds to market equilibrium.

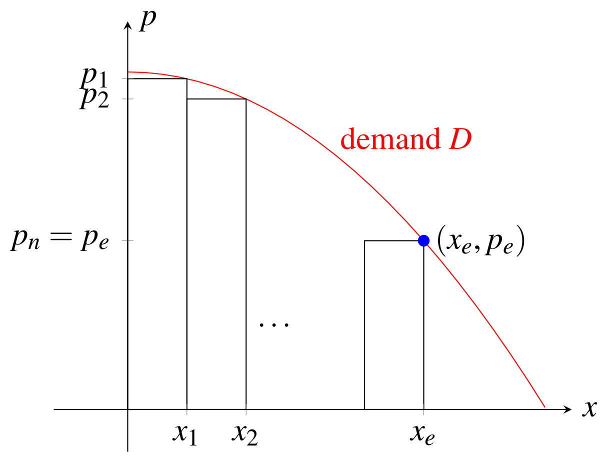
In a competitive market, the price of a commodity will eventually settle at the market equilibrium, which occurs when the supply of the commodity will be equal to its demand as indicated with the point (x_e, p_e) in Figure 3.5. Let us look at a certain commodity, say, dairy to discuss the economic significance of the market equilibrium. As long as $p < p_e$, then the demand for dairy exceeds its supply (see Figure 3.5), which pushes up the price until it reaches the equilibrium price p_e, which in turn signifies that the quantity supplied is equal to the quantity demanded, namely x_e. On the other hand, if $p > p_e$, then the supply of dairy exceeds demand (see Figure 3.5), which brings the price down. In an ideal free market, buying and selling at the equilibrium price should benefit both consumers and producers. In this section, we will compute the **surplus**, which tells us exactly how much the consumers save and the producers gain by buying and selling respectively at the equilibrium price rather than at a higher price.

We begin by computing exactly how much consumers spent when they buy at the equilibrium price p_e:

$$\begin{aligned}
&\text{total amount spent at equilibrium price} \\
&= (\text{number of units bought at equilibrium price}) \cdot (\text{unit price}) \\
&= x_e \cdot p_e
\end{aligned}$$

The quantity $x_e \cdot p_e$ is the rectangular area shown in Figure 3.6a.

Now we compute the total amount that would be spent if every consumer paid the maximum price that each is willing to pay. Given a demand function D, partition the interval $[0, x_e]$ into n subintervals of equal width $\Delta x = \frac{x_e}{n}$ with endpoints $x_i = \frac{ix_e}{n}$, $i = 0, 1, \ldots, n$ as shown below:



Now, let us analyze this partition subinterval by subinterval. Suppose that only x_1 units had been available, then the maximum unit price could have been set at $D(x_1)$ dollars and a total of x_1 units would be sold, but at this price no further units would have been sold. Then the total expenditure in dollars is given by

$$\text{(price per unit)} \cdot \text{(number of units)} = D(x_1) \cdot \Delta x.$$

Now suppose that more units become available by producing x_2 units of our commodity. If the maximum price is set at $D(x_2)$ dollars, then the remaining $x_2 - x_1 = \Delta x$ units can be sold at a cost of $D(x_2) \cdot \Delta x$ dollars. If we continue with this process of price discrimination, then the total amount spent at maximum price is approximately equal to

$$D(x_1) \cdot \Delta x + D(x_2) \cdot \Delta x + \cdots + D(x_n) \cdot \Delta x.$$

Note:

1. On each of the subintervals $[0, x_1]$, $[x_1, x_2]$, …, $[x_{n-1}, x_e]$, the buyers paid as much for each unit as it was worth to them.
2. We recognize that the last sum is a Riemann sum, which yields $\int_0^{x_e} D(x)\,dx$ as $n \to \infty$ or alternatively as $\Delta x \to 0$.

Of course, we simply have calculated the area under the demand curve on the interval $[0, x_e]$, which is shown in Figure 3.6b.

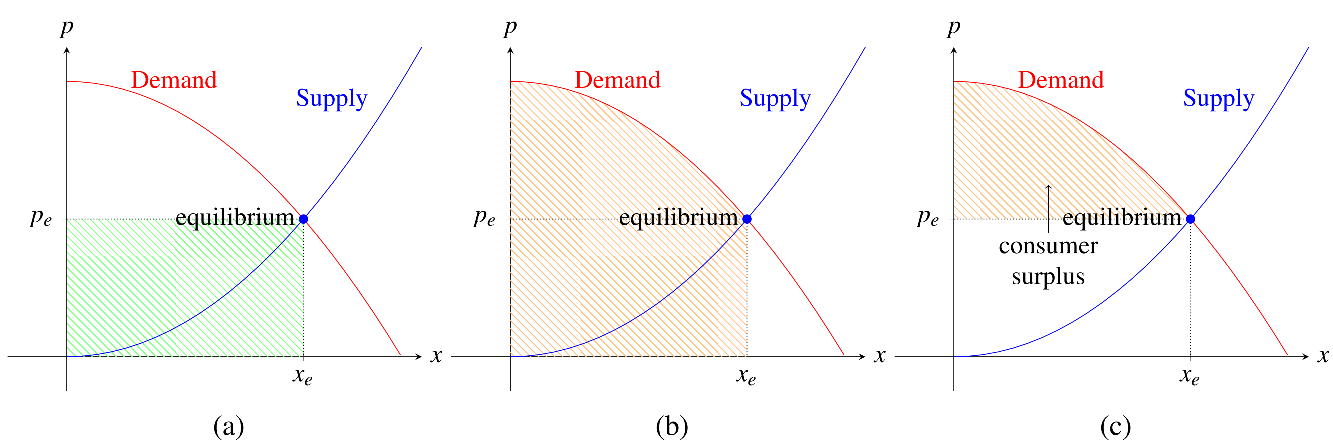


Figure 3.6: Consumer expenditure on $[0, x_e]$ at equilibrium price, in total, and for surplus.

The difference between the total amount spent at the maximum price (Figure 3.6b) and the consumer expenditure at the equilibrium price (Figure 3.6a) is the total amount that consumers save by buying at the equilibrium price (Figure 3.6c). This saving is called the **consumer surplus** for this product. We summarize our result.

Theorem 3.10: Consumer Surplus

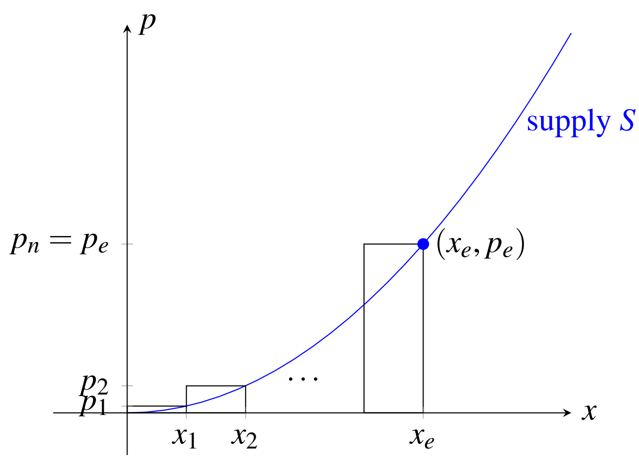
Let D be the demand function, p_e be the equilibrium price, and x_e be the equilibrium quantity sold. Then the **consumer surplus** *is*

$$\int_0^{x_e} D(x)\,dx - x_e p_e = \int_0^{x_e} [D(x) - p_e]\,dx.$$

In a similar manner, we can determine how much producers gain when they sell at the equilibrium price p_e:

$$\begin{aligned}&\text{total amount gained at equilibrium price}\\ &= (\text{number of units sold at equilibrium price}) \cdot (\text{unit price})\\ &= x_e \cdot p_e\end{aligned}$$

Now we compute the total amount that would be gained if every producer sold at the minimum amount they are willing to accept for the product. Given a supply function S, partition the interval $[0, x_e]$ into n subintervals of equal width $\Delta x = \frac{x_e}{n}$ with endpoints $x_i = \frac{ix_e}{n}$, $i = 0, 1, \ldots, n$ as shown below:



Similarly to the subinterval by subinterval analysis for the demand curve, an analysis of the partition of the supply curve shows that the total revenue on every subinterval in dollars by selling at the minimum price is given by

$$(\text{price per unit}) \cdot (\text{number of units}) = S(x_i) \cdot \Delta x$$

for $i = 1, 2, \ldots, n$. Continuing with the process of price discrimination, the total amount gained at minimum price is approximately equal to

$$S(x_1) \cdot \Delta x + S(x_2) \cdot \Delta x + \cdots + S(x_n) \cdot \Delta x.$$

Note:

1. On each of the subintervals $[0, x_1]$, $[x_1, x_2], \ldots [x_{n-1}, x_e]$, the producers sold at the lowest price that they are willing to set.
2. This sum is again a Riemann sum, which yields $\int_0^{x_e} S(x)\,dx$ as $n \to \infty$ or alternatively as $\Delta x \to 0$.

Figure 3.7b shows the area under the supply curve on the interval $[0, x_e]$ that our computation has yielded.

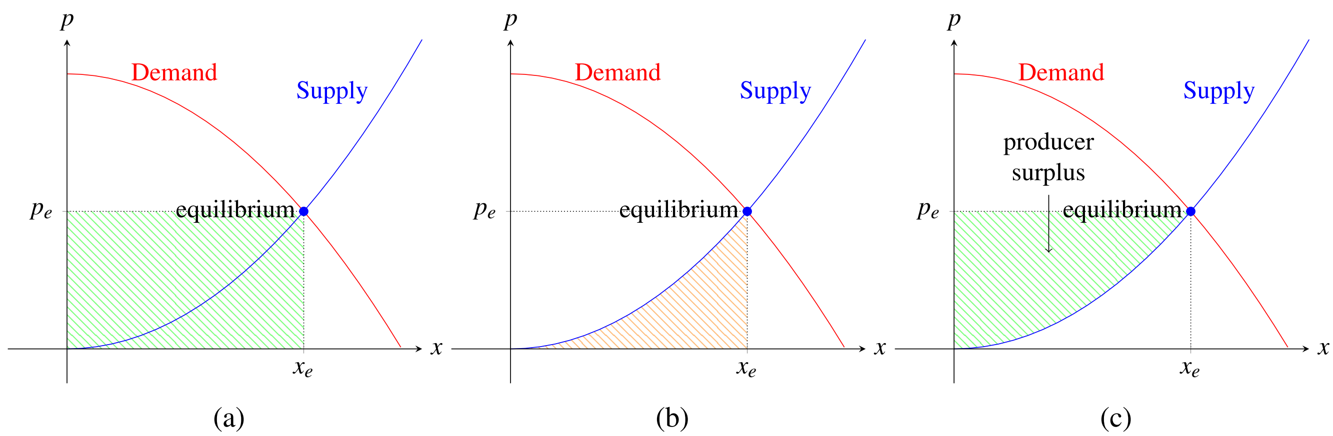


Figure 3.7: Producer revenue on $[0, x_e]$ at equilibrium price, in total, and for surplus.

The difference between the producer revenue at the equilibrium price (Figure 3.7a) and the total amount achieved at the minimum price (Figure 3.7b) is the total amount that producers gain by selling at the equilibrium price (Figure 3.7c). This income is called the **producer surplus** for this product. We summarize our result.

Theorem 3.11: Producer Surplus

Let S be the supply function, p_e be the equilibrium price, and x_e be the equilibrium quantity sold. Then the ***producer surplus*** *is*

$$x_e p_e - \int_0^{x_e} S(x)\,dx = \int_0^{x_e} [p_e - S(x)]\,dx.$$

Note: In general, the consumer surplus and the producer surplus are not equal.

Example 3.12: Consumer Surplus

The demand for a product, in dollars, is

$$p = D(x) = 1000 - 0.5x - 0.0002x^2.$$

Find the consumer surplus when the sales level is 200.

Solution. When the number of units sold is $x_e = 200$, the corresponding price is

$$p_e = 1000 - 0.5(200) - 0.0002(200^2) = 892.$$

Therefore, the consumer surplus is

$$\begin{aligned}
\int_0^{200} [D(x) - p_e]\, dx &= \int_0^{200} \left(1000 - 0.5x - 0.0002x^2 - 892\right) dx \\
&= \int_0^{200} \left(108 - 0.5x - 0.0002x^2\right) dx \\
&= \left[108x - 0.25x^2 - \frac{0.0002}{3}x^3\right]_0^{200} \\
&= 108(200) - 0.25(200)^2 - \frac{0.0002}{3}(200)^3 \\
&= \$11,066.70
\end{aligned}$$

♣

Example 3.13: Consumer and Producer Surplus

The price, in dollars per kg, of flour at a certain grocer is given by

$$D(x) = 75 - e^{0.2x}, \qquad \text{for } x \in [0,21]$$

corresponding to a demand of x kg. When the supply is x kg, the wholesale price is given by

$$S(x) = 2e^{0.2x} - 1.$$

Determine the consumer surplus and the producer surplus.

Solution. We set $D(x) = S(x)$ to find the equilibrium quantity and price:

$$\begin{aligned}
75 - e^{0.2x} &= 2e^{0.2x} - 1 \\
e^{0.2x} &= \frac{76}{3} \\
x &= \frac{\ln\left(\frac{76}{3}\right)}{0.2} \approx 16.16 \text{ kg.}
\end{aligned}$$

Therefore, the equilibrium point is

$$(x_e, p_e) = (16.16, 75 - e^{0.2(16.16)}) \approx (16.16, 49.67).$$

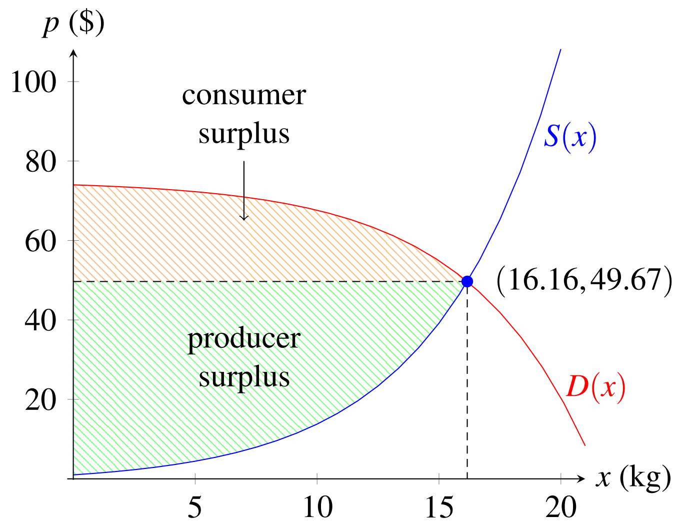
We now calculate the consumer surplus as

$$\begin{aligned}
\int_0^{16.16} \left[(75 - e^{0.2x}) - 49.67\right] dx &= \int_0^{16.16} \left(25.33 - e^{0.2x}\right) dx \\
&= \left[25.33x - \frac{1}{0.2}e^{0.2x}\right]_0^{16.16} \\
&\approx \$287.68.
\end{aligned}$$

Similarly, we calculate the producer surplus as

$$\begin{aligned}\int_0^{16.16} \left[49.67 - \left(2e^{0.2x} - 1\right)\right]\, dx &= \int_0^{16.16} \left(50.67 - 2e^{0.2x}\right)\, dx \\ &= \left[50.67x - \frac{2}{0.2}e^{0.2x}\right]_0^{16.16} \\ &\approx \$575.53.\end{aligned}$$

Thus, the consumer surplus is approximately \$287.68 and the producer surplus is approximately \$575.53. These quantities are illustrated in the figure below.

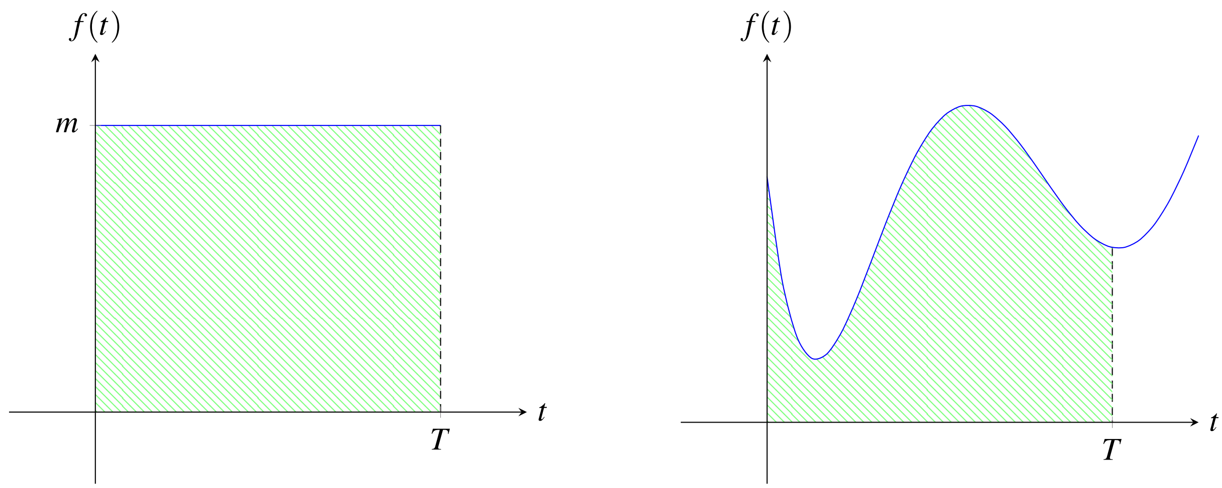


♣

3.2.2. Continuous Money Flow

Another application of integrals is for the computation involved in **continuous money flow** problems. These problems may be payments to repay a debt or deposits to establish a retirement plan that are made at regular monthly payments over a certain time period. If the time between deposits or payments are made in relatively short intervals compared to the entire time period, then the money can be thought of as flowing continuously, hence the name *continuous* money flow. Some textbooks refer to these types of problems as "continuous income streams" or "continuous money streams". Naturally, deposits or payments cannot be made continuously in a physical manner, since the units of legal tender are by their very nature discrete such as the cent and dollar.

A word of caution: Discussing the value of money is difficult as there are many parameters affecting it. For the purpose of this introductory section, we take a simplified point of view such as assuming that there is a single interest rate, no inflation, and no other factors influencing computation. This simplified approach will provide for an intuitive understanding of solving continuous money flow problems that can then be expanded upon. Before we develop formulas for the present and future value of such a continuous money flow, we begin by computing the total amount of money that is being accumulated making the regular deposits or payments.



(a) *A uniform rate of money flow $f(t) = m$ dollars per year.* (b) *A variable rate of money flow $f(t)$ dollars per year.*

Figure 3.8

Suppose that $f(t)$ is the rate in dollars per unit of time for a continuous money flow. Figure 3.8 above shows a uniform rate of money flow $f(t) = m$ dollars per year to the left and a variable rate of money flow $f(t)$ dollars per year to the right. The **total income** in dollars of the former case can be readily calculated and is simply the rectangular area $m \cdot T$ dollars. For the latter case, we readily understand that the area under the curve f from $t = 0$ to $t = T$ is the total income, whose value in dollars is determined by the definite integral $\int_0^T f(t)\,dt$.

Theorem 3.14: Total Income of Money Flow

Suppose that $f(t)$ is the rate in dollars per unit of time for a continuous money flow. Then the **total income** I *over the time period $[0,T]$ is*

$$I = \int_0^T f(t)\,dt.$$

Note: The above computation does not calculate the accumulated interest, just the total income.

Example 3.15: Total Income

Suppose the income of a small coffee shop projected to grow exponentially. The rate of income when the shop first opened was at \$5,000 per month, which grew to a rate of \$6,000 per month by the second month. What is the expected total income from the coffee shop during its first 6 months of operation?

Solution. Let t be the time in months since the opening of the shop. The projected income is of the form

$$f(t) = Ce^{kt}.$$

Since $f(0) = 5000$, we must have $C = 5000$. Now to determine k, we use $f(1) = 6000$:

$$\begin{aligned} 5000e^{k(1)} &= 6000 \\ k &= \ln \frac{6000}{5000} \\ k &\approx 0.18. \end{aligned}$$

Thus, $f(t) = 5000e^{0.18t}$. We can now determine the total income, I:

$$\begin{aligned} I &= \int_0^6 f(t)\,dt \\ &= \int_0^6 5000e^{0.18t}\,dt \\ &= \frac{5000}{0.18} e^{0.18t} \Big|_0^6 \\ &\approx 54018.8764. \end{aligned}$$

We conclude that the coffee shop is expected to produce an income of approximately \$54,018.88 in its first six months of operation. ♣

Let us now develop the formulas for the present value and future value of a continuous money flow. We begin by summarizing the present and future value of money that earns interest compounded continuously at a rate r over time t without making continuous deposits or payments. Suppose that F is the amount of money at time t that accumulates interest compounded continuously at a rate r over time t, then the growth of the investment or loan due to accumulated interest earned over time is described by the differential equation

$$\frac{dF}{dt} = rF.$$

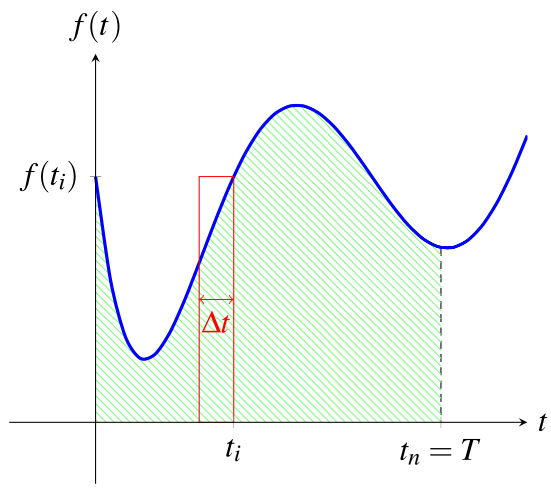
In Chapter 5, differential equations will be discussed in more detail, but for now, we simply present the solution to this differential equation, and so the *future value* is described by

$$F(t) = Pe^{rt},$$

where P is the initial amount. Similarly, if we want to obtain an amount of F dollars that has accrued interest compounded continuously at the annual rate r over t years, then the amount of money that needs to be deposited today, namely the present value P, is given by

$$P(t) = Fe^{-rt}.$$

Now, the **present value** P of a continuous money stream must be the amount of money that will accumulate to the same future value as the continuous money when deposited at the same interest rate compounded continuously for the same period of time but without the continuous stream of income. Let $f(t)$ represent the rate of continuous money flow and r the rate of interest for T years. As usual, we partition the interval $[0, T]$ into n subintervals of equal width $\Delta t = \frac{T}{n}$ with endpoints $t_i = \frac{iT}{n}$, $i = 0, 1, \ldots, n$ as shown below.



Using our result from above about accruing interest compounded continuously, the present value P_i of the money flow over the i-th subinterval is approximately

$$P_i = [f(t_i)\Delta t]\, e^{-rt_i},$$

and so the total present value is approximated by a typical Riemann sum

$$P \approx \sum_{i=1}^{n} [f(t_i)\Delta t]\, e^{-rt_i}.$$

Taking the limit as $n \to \infty$, we obtain the following definite integral

$$P = \int_0^T f(t)e^{-rt}\, dt.$$

We capture our result in the following theorem.

Theorem 3.16: Present Value of a Continuous Money Flow

Suppose that $f(t)$ is the rate of money flow in dollars per unit of time that is earning interest compounded continuously at a rate r per unit of time over the time period $[0, T]$. Then the **present value** P *of this continuous money flow is*

$$P = \int_0^T f(t)e^{-rt}\, dt.$$

Before we develop the formula for the accumulated amount of a continuous money flow, we present one example of a present value problem.

Example 3.17: Present Value

A certain manufacturer produces industrial bread-making machines, which are estimated to generate a continuous income stream each with a rate of $28 - 4.5t$ million dollars per year in year t of operation. Each machine is in operation for about 6 years and the money is invested at the annual rate of 5% compounded continuously. Compute the fair market price of each machine.

Solution. We assume that the fair market price of each machine is the present value of the continuous money flow with

$$f(t) = 28 - 4.5t, \ r = 0.05, \text{ and } T = 6.$$

Therefore, the price P is given by

$$P = \int_0^6 [28 - 4.5t]\, e^{-0.05t}\, dt = \int_0^6 28e^{-0.05t}\, dt - 4.5 \int_0^6 te^{-0.05t}\, dt.$$

The first integral can be computed directly

$$\int_0^6 28e^{-0.05t}\, dt = \left[-\frac{28}{0.05}e^{-0.05t}\right]_0^6 = 560\left(1 - e^{-0.03}\right).$$

The second integral can be solved using Integration by Parts, with

$$u = t, \ dv = e^{-0.05t} dt \implies du = dt, \ v = -\frac{1}{0.05}e^{-0.05t}$$

to obtain

$$\begin{aligned}\int_0^6 te^{-0.05t}\, dt &= \left[-\frac{t}{0.05}e^{-0.05t}\right]_0^6 + \frac{1}{0.05}\int_0^6 e^{-0.05t}\, dt \\ &= -\left[\frac{0.05t + 1}{0.05}e^{-0.05t}\right]_0^6 = 20 - 26e^{-0.3}.\end{aligned}$$

Hence,

$$\begin{aligned}P &= \int_0^6 [28 - 4.5t]\, e^{-0.05t}\, dt \\ &= 560\left(1 - e^{-0.3}\right) - 4.5\left(20 - 26e^{-0.3}\right) \\ &\approx 141.82,\end{aligned}$$

and so the fair market price of each machine is about 142 million dollars. ♣

Similarly, we can develop a formula for the **future value** of a continuous money flow. Again, we partition the interval $[0, T]$ into n subintervals of equal width $\Delta t = \frac{T}{n}$ with endpoints $t_i = \frac{iT}{n}$, $i = 0, 1, \ldots, n$. Now, the money that is deposited or paid during the i-th time interval stays in the account for the length of time $T - t_i$. Using our result from the beginning of this section about interest being earned compounded continuously, the future value F_i of the money flow over the i-th subinterval is approximately

$$F_i = [f(t_i)\Delta t]\, e^{r(T - t_i)},$$

and so the total future value is again approximated by a Riemann sum, namely

$$F \approx \sum_{i=1}^{n} [f(t_i)\Delta t]\, e^{r(T - t_i)}.$$

Taking the limit as $n \to \infty$, the following definite integral is obtained

$$F = \int_0^T f(t)e^{r(T-t)}\, dt.$$

This leads to the following theorem.

Theorem 3.18: Future Value of a Continuous Money Flow

Suppose that $f(t)$ is the rate of money flow in dollars per unit of time that is earning interest compounded continuously at a rate r per unit of time over the time period $[0,T]$. Then the **future value** *F of this continuous money flow is*

$$F = \int_0^T f(t)e^{r(T-t)}\,dt.$$

Example 3.19: Continuous Money Flow

Suppose money is flowing continuously at a constant rate $f(t)$ for 10 years at 8% interest compounded continuously accumulating to \$20,000. Compute $f(t)$.

Solution. We have that

$$F = 20,000, \ r = 0.08, \ \text{and } T = 10.$$

We can use the formula for the future value to solve for the rate $f(t)$ of the continuous money flow:

$$20,000 = \int_0^{10} f(t)e^{0.08(10-t)}\,dt$$

Since the rate of money flow is constant, $f(t)$ does not depend on t. Therefore, $f(t)$ can be taken out of the integral as follows.

$$\begin{aligned} 20,000 &= f(t)\int_0^{10} e^{0.08(10-t)}\,dt \\ -0.08\cdot e^{-0.8}\cdot 20,000 &= f(t)\left[e^{-0.08t}\right]_0^{10} \\ -0.08\cdot e^{-0.8}\cdot 20,000 &= f(t)\left[e^{-0.8}-e^{0}\right] \\ f(t) &= \frac{0.08\cdot 20,000}{[e^{0.8}-1]} \approx \$1305.55 \end{aligned}$$

♣

Example 3.20: Total Income, Present and Future Values

Suppose money is flowing continuously at a constant rate of \$3500 per year for 10 years at 1.25% interest compounded continuously. Compute the following:

(a) total income;

(b) future value at $T = 10$ years;

(c) total interest; and

(d) present value.

Solution. We have that $f(t) = 3500$, $r = 0.0125$, and $T = 10$.

(a) The total income I is given by

$$I = \int_0^{10} 3500\,dt = \$35,000.$$

(b) The future value F at $T = 10$ years is calculated to be

$$F = \int_0^{10} 3500e^{0.0125(10-t)}\,dt = -\frac{3500}{0.0125}\left[e^{0.0125(10-t)}\right]_0^{10} = \$37,281.75.$$

(c) The total interest is the difference between the future value F at $T = 10$ years and the total income over 10 years:

$$\$37,281.75 - \$35,000 = \$2281.57.$$

(d) The present value P is obtained as

$$\begin{aligned} P &= \int_0^{10} 3500e^{-0.0125t}\,dt \\ &= \left[-\frac{3500}{0.0125}e^{-0.0125}\right]_0^{10} \\ &= \frac{3500}{0.0125}\left(1 - e^{-0.125}\right) \approx \$39,900.87. \end{aligned}$$

♣

Exercises for Section 3.2

Exercise 3.2.1 *Suppose the supply function for a certain product is given by*

$$S(q) = q^{5/2} + 3q^{1/2} + 25.$$

If the equilibrium quantity is $q = 10$, determine the producer surplus.

Exercise 3.2.2 *It is determined that the supply function for a certain nutritional supplement can be described by*

$$S(x) = x^{3/2} + 2x^{1/2} + 50.$$

Assuming that the supply and demand are in equilibrium when $x = 12$, determine the producer surplus.

Exercise 3.2.3 *Determine the consumer surplus corresponding to the demand function*

$$D(q) = \frac{100}{(q+2)^2},$$

if the equilibrium quantity is $q = 2$.

Exercise 3.2.4 *Suppose the demand function for organic tomatoes can be described by*

$$D(x) = \frac{20,000}{(2x+5)^3}.$$

Determine the consumer surplus if the supply and demand are in equilibrium at $x = 4$.

Exercise 3.2.5 *The supply function for steel-cut oats is found to be*

$$S(q) = q(2q+25),$$

dollars and the demand function is found to be

$$D(q) = 1000 - 10q - q^2,$$

dollars, where q is measured in kgs.

(a) Determine the equilibrium point.

(b) What is the consumer surplus?

(c) What is the producer surplus?

(d) Display your results on a graph of the supply and demand curves.

Exercise 3.2.6 *It is determined that the supply function for a certain product x (measured in units of thousands) can be described by*

$$S(x) = \left(2x + \frac{1}{4}\right)^2,$$

with a corresponding demand function

$$D(x) = \frac{200}{15x+1}.$$

(a) Determine the equilibrium point.

(b) What is the consumer surplus?

(c) What is the producer surplus?

(d) Display your results on a graph of the supply and demand curves.

Exercise 3.2.7 *For each of the following rate of money flow in dollars per year, rate r of continuously compounded interest, and time period T years, compute (i) the total income; (ii) the future value at T; (iii) the total interest; and (iv) the present value.*

(a) $f(t) = 200,\ r = 2\%,\ T = 5.$

(b) $f(t) = 4000,\ r = 0.5\%,\ T = 7.$

(c) $f(t) = 10,000,\ r = 1\%, T = 4.$

(d) $f(t) = 300t,\ r = 1\%, T = 4.$

(e) $f(t) = 450e^{0.5t},\ r = 1.5\%, T = 15.$

(f) $f(t) = 100t + 50,\ r = 3\%,\ T = 9.$

Exercise 3.2.8 *The value of a continuous money flow is found to be \$5,000 after 6 years. Assume the rate of money flow $f(t)$ is constant, and the interest is continuously compounded at a rate of 1%. Compute $f(t)$.*

Exercise 3.2.9 *The value of a continuous money flow is found to be \$10,000 after 5 years, and \$14,000 after 7 years. The rate of money flow $f(t)$ is linear, and the interest is compounded continuously at a rate of 1.5%. Compute $f(t)$.*

3.3 Volume of Revolution: Disk Method

We have seen how to compute certain areas by using integration; we will now look into how some volumes may also be computed by evaluating an integral. Generally, the volumes that we can compute this way have cross-sections that are easy to describe. For example, circular cross-sections are easy to describe as their area just depends on the radius, and so they are one of the central topics in this section. However, we first discuss the general idea of calculating the **volume** of a solid by slicing up the solid.

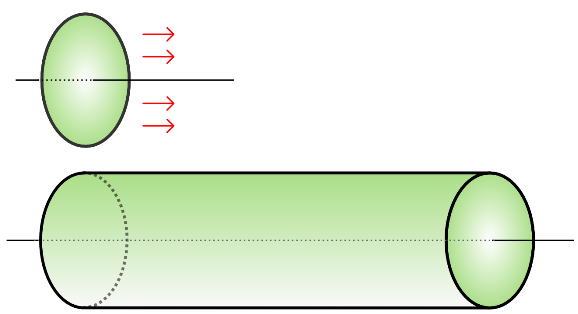
3.3.1. Computing Volumes with Cross-sections

Let us first formalize what is meant by a **cross-section**.

Definition 3.21: Cross-section

A **cross-section** *of a solid is the region obtained by intersecting the solid with a plane.*

Examples of cross-sections are the circular region above the right cylinder in Figure 3.9a, the star above the star-prism in Figure 3.9b, and the square we see in the pyramid on the left side of Figure 3.10. For now, we are only interested in solids, whose volumes are generated through cross-sections that are easy to describe. For example, the right cylinder in Figure 3.9a is generated by translating a circular region along the x-axis for a certain length h. Every cross-section of the right cylinder must therefore be circular, when cutting the right cylinder anywhere along length h that is perpendicular to the x-axis.


(a) *Translated disk.*

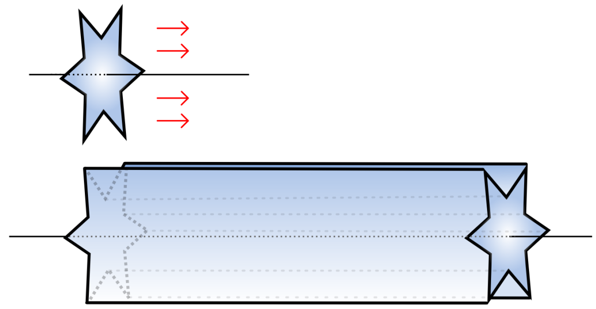

(b) *Translated star.*

Figure 3.9

In a similar manner, many other solids can be generated and understood as shown with the translated star in Figure 3.9b.

The volume of both the right cylinder and the translated star can be thought of as

$$V = (\text{area of cross-section}) \cdot (\text{length}) = A \cdot h.$$

Let us now turn towards the calculation of such volumes by working through two examples. The right pyramid with square base shown in Figure 3.10 has cross-sections that must be squares if we cut the pyramid parallel to its base. The following example makes use of these cross-sections to calculate the volume of the pyramid for a certain height.

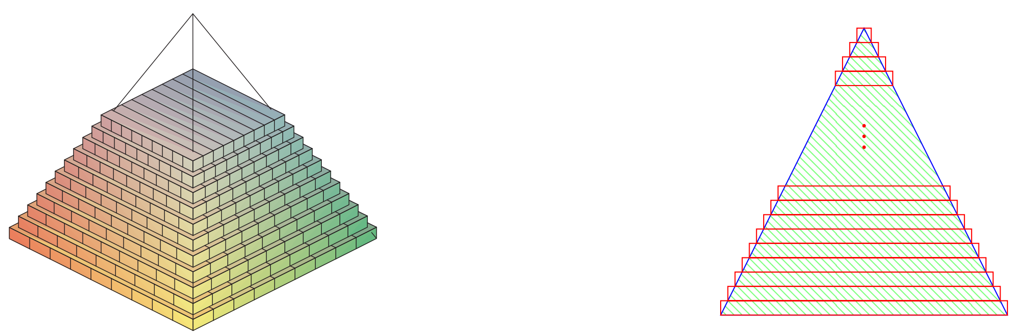


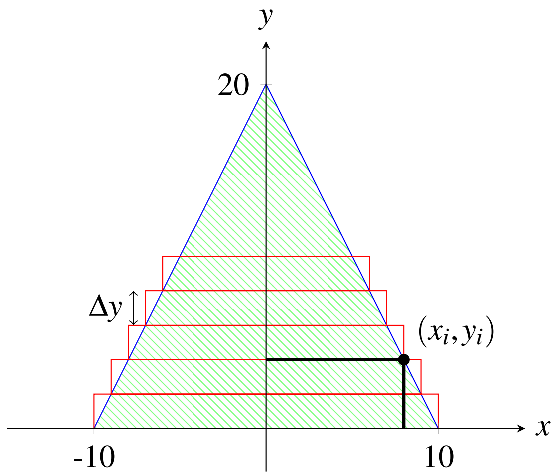
Figure 3.10: Volume of a pyramid approximated by rectangular prisms.

Example 3.22: Volume of a Pyramid

Find the volume of a pyramid that is 20 metres tall with a square base 20 metres on a side.

Solution. As with most of our applications of integration, we begin by asking how we might approximate the volume. Since we can easily compute the volume of a rectangular prism (that is, a "box"), we will use some boxes to approximate the volume of the pyramid, as shown in Figure 3.10: Suppose we cut up the pyramid into n slices. On the left is a 3D view that shows cross-sections cut parallel to the base of the pyramid and replaced with rectangular boxes that are used to approximate the volume. On the right is a 2D view that now shows a cross-section perpendicular to the base of the pyramid so that we can identify the width and height of a box.

Having to use width and height means that we have two variables. However, by overlaying a Cartesian coordinate system with the origin at the midpoint of the base on to the 2D view of Figure 3.10 as shown below, we can relate these two variables to each other.



Since the cross-sectional view is placed symmetrically about the y-axis, we see that a height of 20 is achieved at the midpoint of the base. Thus at $x = 0$, $y = 20$, at $x = 10$, $y = 0$, and we have a slope of $m = -2$. So

$$\begin{aligned} y &= -2x + b \\ 20 &= -2(0) + b \\ 20 &= b. \end{aligned}$$

Therefore, $y = 20 - 2x$, and in the terms of x we have that $x = 10 - y/2$. We could have also used similar triangles here to derive the relationship between x and y. From the right diagram in Figure 3.10, we see that each box has volume of the form

$$(2x_i)(2x_i)\Delta y.$$

Then the total volume is approximately

$$\sum_{i=0}^{n-1}(2x_i)(2x_i)\Delta y = \sum_{i=0}^{n-1} 4(10 - \frac{y_i}{2})^2 \Delta y$$

and when we apply the limit $\Delta y \to 0$ we get the volume as the value of a definite integral as defined in Section 1.4:

$$\begin{aligned} V &= \lim_{\Delta y \to 0} \sum_{i=0}^{n-1} 4(10 - \frac{y_i}{2})^2 \Delta y = \int_0^{20} 4(10 - \frac{y}{2})^2 \, dy \\ &= \int_0^{20} (20 - y)^2 \, dy \\ &= -\frac{(20-y)^3}{3}\Big|_0^{20} \\ &= -\frac{0^3}{3} - \left(-\frac{20^3}{3}\right) = \frac{8000}{3}. \end{aligned}$$

As you may know, the volume of a pyramid is given by the formula

$$(1/3)(\text{height})(\text{area of base}).$$

Using this formula we calculate

$$(1/3)(20)(400) = \frac{8000}{3},$$

which agrees with our answer. ♣

Example 3.23: Volume of an Object

The base of a solid is the region between $f(x) = x^2 - 1$ and $g(x) = -x^2 + 1$ as shown to the right of Figure 3.11, and its cross-sections perpendicular to the x-axis are equilateral triangles, as indicated in Figure 3.11 to the left.
The solid has been truncated to show a triangular cross-section above $x = 1/2$.
Find the volume of the solid.

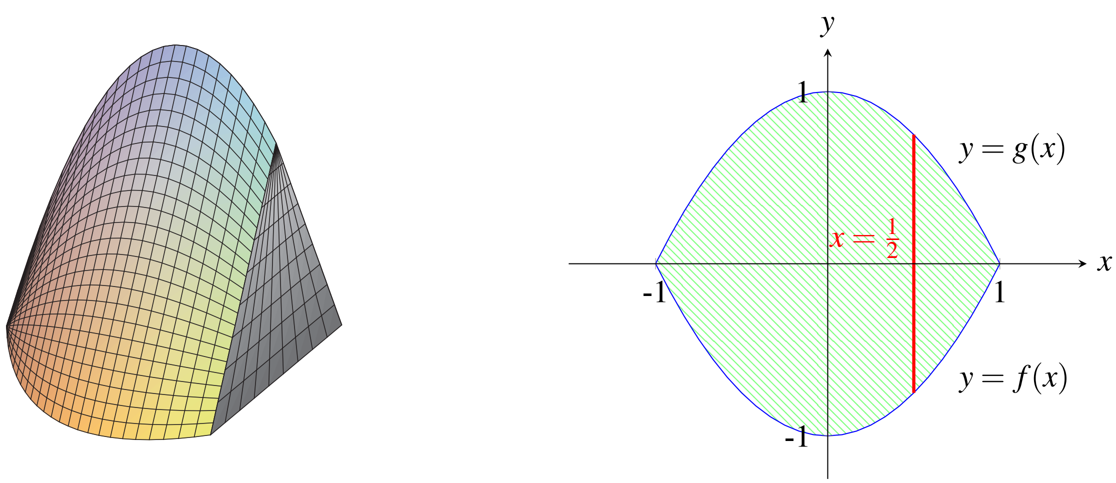
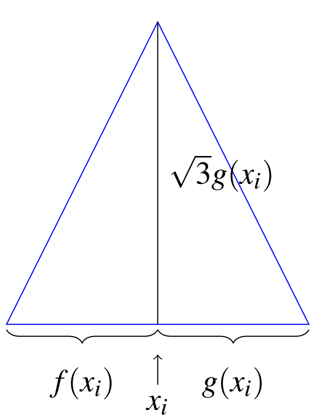
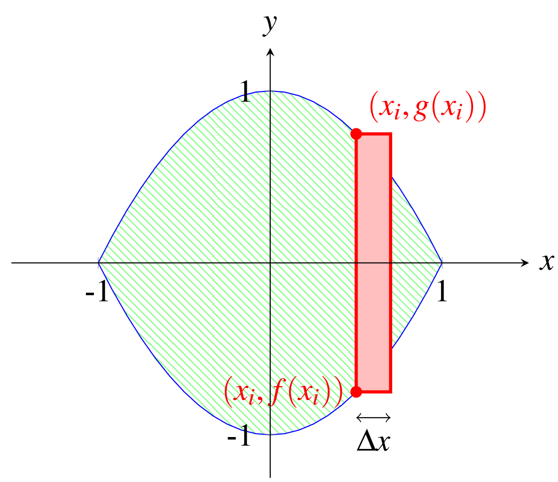


Figure 3.11: Solid with equilateral triangles as cross-sections.

Solution. We begin by drawing the equilateral triangle above any x_i and identify its base and height as shown below to the left.





The equilateral triangle has base

$$g(x_i) - f(x_i) = (1 - x_i^2) - (x_i^2 - 1) = 2(1 - x_i^2),$$

and height

$$\sqrt{3} g(x_i) = \sqrt{3}(1 - x_i^2).$$

So the area of the cross-section is

$$\frac{1}{2}(\text{base})(\text{height}) = (1 - x_i^2)\sqrt{3}(1 - x_i^2).$$

The diagram above to the right indicates the position of an arbitrary thin equilateral triangle in the given region. Therefore, the volume of this thin equilateral triangle is given by

$$\frac{1}{2}(\text{base})(\text{height})(\text{thickness}) = (1 - x_i^2)\sqrt{3}(1 - x_i^2)\Delta x.$$

If we have sliced our solid into n thin equilateral triangles, then the volume can be approximated with the sum

$$\sum_{i=0}^{n-1} (1 - x_i^2)\sqrt{3}(1 - x_i^2)\Delta x = \sum_{i=0}^{n-1} \sqrt{3}(1 - x_i^2)^2 \Delta x.$$

Similar to the previous example, when we apply the limit $\Delta x \to 0$, the total volume is

$$V = \lim_{\Delta x \to 0} \sum_{i=0}^{n-1} \sqrt{3}(1 - x_i^2)^2 \Delta x = \int_{-1}^{1} \sqrt{3}(1 - x^2)^2 \, dx = \frac{16}{15}\sqrt{3}.$$

Notice that the limits of integration, namely -1 and 1, are the left and right bounding values of x, because we are slicing the solid perpendicular to the x-axis from left to right. ♣

3.3.2. Disk Method: Integration w.r.t. x

One easy way to get "nice" cross-sections is by rotating a plane figure around a line, also called the **axis of rotation**, and therefore such a solid is also referred to as a **solid of revolution**. For example, in Figure 3.12 we see a plane region under a curve and between two vertical lines $x = a$ and $x = b$, which creates a solid when the region is rotated about the x-axis, and naturally, a typical cross-section perpendicular to the x-axis must be circular as shown.

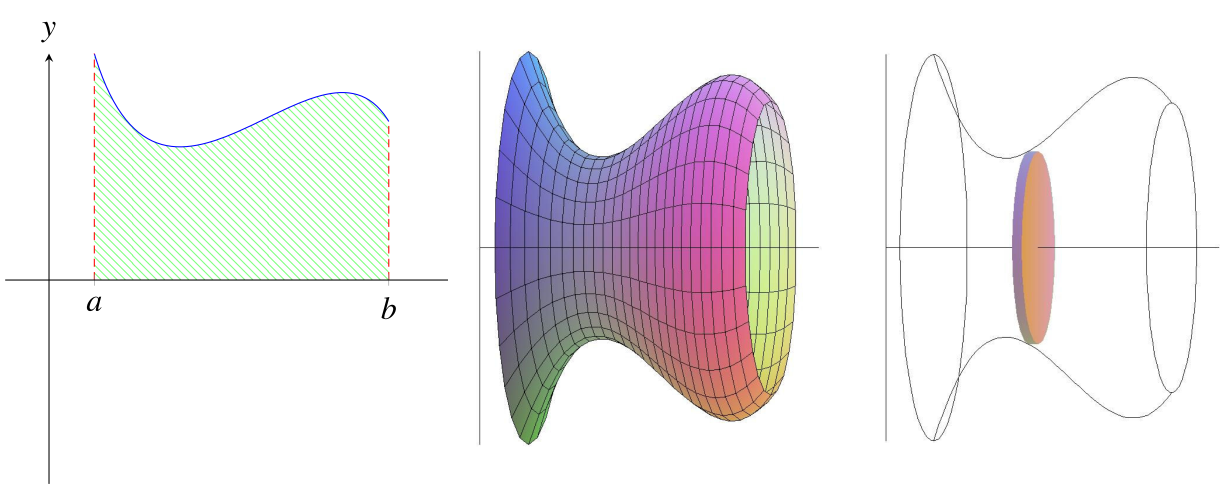


Figure 3.12: A solid of rotation.

Of course a real "slice" of this figure will not be cylindrical in nature, but we can approximate the volume of the slice by a cylinder or so-called **disk** with circular top and bottom and straight sides parallel to the axis of rotation; the volume of this disk will have the form $\pi r^2 \Delta x$, where r is the radius of the disk and Δx is the thickness of the disk. As long as we can write r in terms of x we can compute the volume by an integral. Often, the radius r is given by the height of the function, i.e. $r = f(x_i)$ and so we compute the volume in a similar manner as in Section 3.3.1:

Suppose there are n disks on the interval $[a,b]$, then the volume of the solid of revolution is approximated by

$$\sum_{i=0}^{n-1} \pi \left[f(x_i)\right]^2 \Delta x,$$

and when we apply the limit $\Delta x \to 0$, the volume computes to the value of a definite integral

$$V = \lim_{\Delta x \to 0} \sum_{i=0}^{n-1} \pi \left[f(x_i)\right]^2 \Delta x = \int_a^b \pi \left[f(x)\right]^2 \, dx.$$

Because the volume of the solid of revolution is calculated using disks, this type of computation is often referred to as the **Disk Method**.

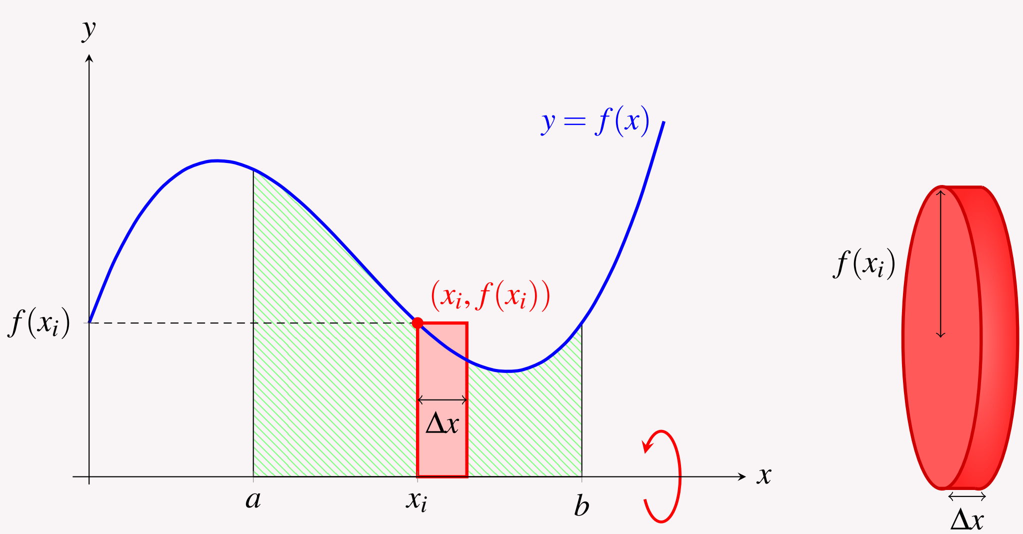
We capture our results in the following theorem.

Theorem 3.24: Disk Method: Integration w.r.t. x

Suppose f is non-negative and continuous on the interval $[a,b]$. Then the volume V formed by rotating the area under the curve of f about the x-axis is

$$V = \lim_{\Delta x \to 0} \sum_{i=0}^{n-1} \pi \left[f(x_i)\right]^2 \Delta x = \int_a^b \pi \left[f(x)\right]^2 \, dx, \text{ where}$$

x_i is the location of the disk,
$f(x_i)$ is the radius of the disk, and
Δx is the thickness of the disk as shown below.



Note: Either of the variations 1 and 2 described below will most likely create disks with holes in the solid of revolution, and so we would have to apply the so-called **Washer Method** that is described in the next section.

1. We could rotate the area of *any* region around an axis of rotation, including the area of a region bounded above by a function $y = f(x)$ and below by a function $y = g(x)$ on an interval $x \in [a,b]$.
2. The axis of rotation can be any axis parallel to the x-axis for this method to work.

We now provide one further example of the Disk Method.

Example 3.25: Volume of a Right Circular Cone

Find the volume of a right circular cone with

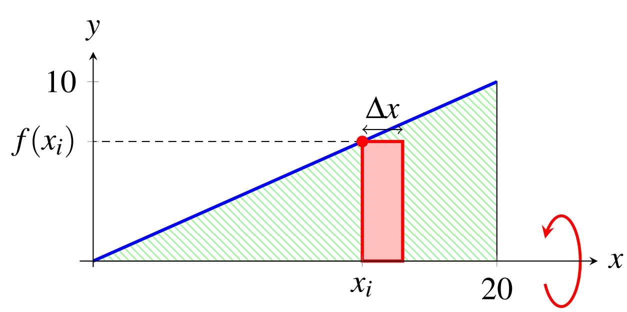
(a) base radius 10 and height 20.

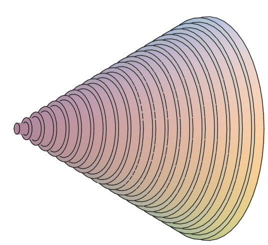
(b) base radius r and height h.

(A right circular cone is one with a circular base and with the tip of the cone directly over the centre of the base.)

Solution.

(a) We can view this cone as produced by the rotation of the line $y = x/2$ rotated about the x-axis, as indicated below.





At a particular point on the x-axis, say x_i, the radius of the resulting cone is the y-coordinate of the corresponding point on the line, namely $y_i = x_i/2$. Thus the total volume is approximately

$$\sum_{i=0}^{n-1} \pi (x_i/2)^2\, dx$$

and the exact volume is

$$\int_0^{20} \pi \frac{x^2}{4}\, dx = \frac{\pi}{4}\frac{x^3}{3}\bigg|_0^{20} = \frac{\pi}{4}\frac{20^3}{3} = \frac{2000\pi}{3}.$$

(b) Note that we can instead do the calculation with a generic height and radius:

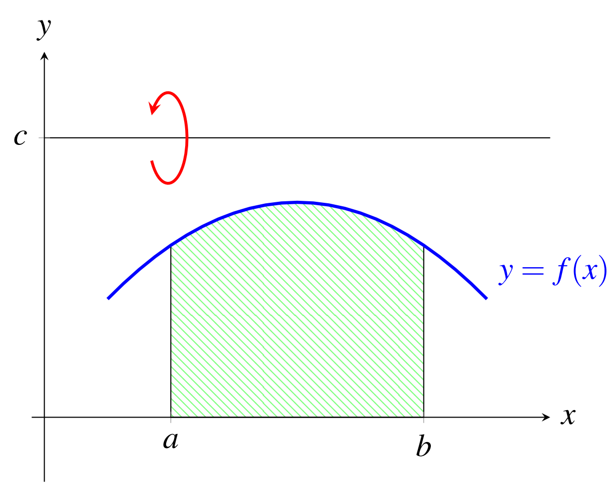
$$\int_0^h \pi \frac{r^2}{h^2} x^2\, dx = \frac{\pi r^2}{h^2}\frac{h^3}{3} = \frac{\pi r^2 h}{3},$$

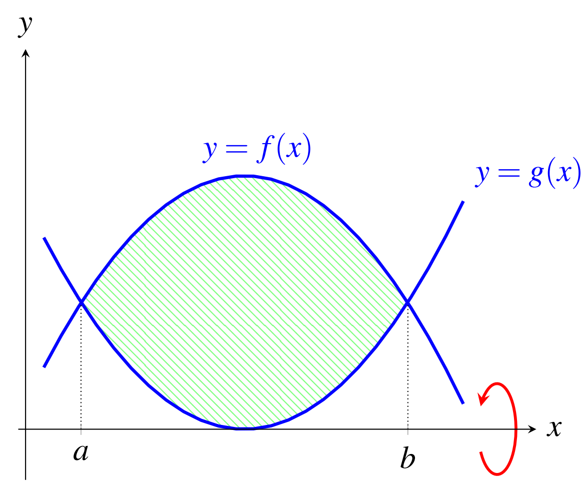
giving us the usual formula for the volume of a cone.

♣

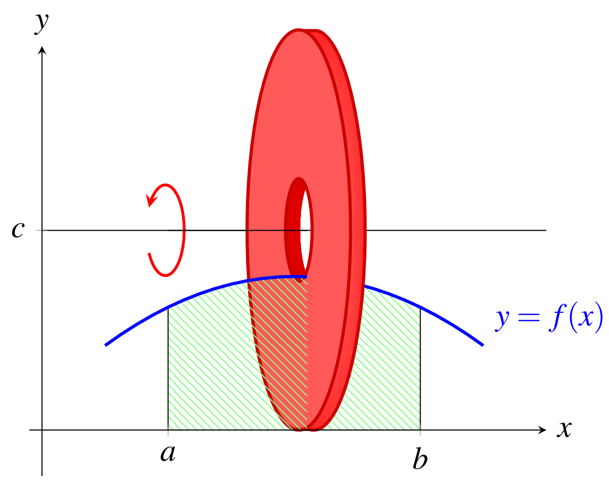
3.3.3. Washer Method: Integration w.r.t x

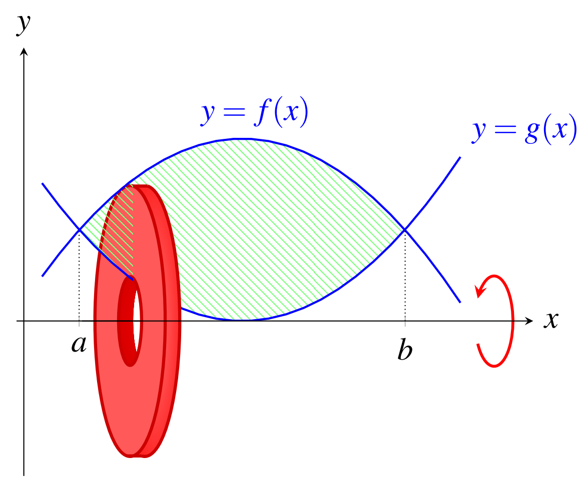
Suppose the axis of revolution is not part of the boundary of an area as shown below in two different scenarios:





When either of the above area is rotated about its axis of rotation, then the solid of revolution that is created has a hole on the inside – like a distorted donut. If we now slice the solid perpendicular to the axis of rotation, then the cross-section shows a disk with a hole in it as indicated below. Such a disk looks like a "washer" and so the method that employs these disks for finding the volume of the solid of revolution is referred to as the **Washer Method**. The following example demonstrates how to find a volume that is created in this fashion.

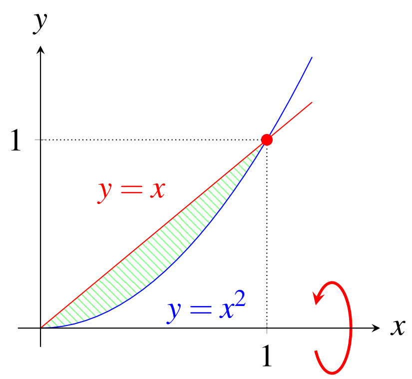


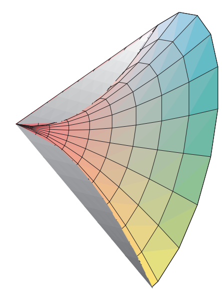


Example 3.26: Volume of an Object with a Hole

Find the volume of the object generated when the area between $y = x^2$ and $y = x$ is rotated around the x-axis.

Solution. We begin by graphing the area between $y = x^2$ and $y = x$ and note that the two curves intersect at the point $(1,1)$ as shown below to the left.

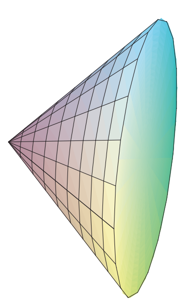


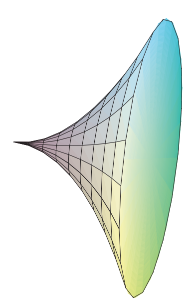


We now rotate this around around the x-axis as shown above to the right. We notice that the solid has a *hole* in the middle and we now consider two methods for calculating the volume.

Method 1: Subtraction of Volumes

We can think of the volume of the solid of revolution as the subtraction of two volumes: the outer volume is that of the solid of revolution created by rotating the line $y = x$ around the x-axis (see left graph in the figure below) – namely the volume of a cone, and the inner volume is that of the solid of revolution created by rotating the parabola $y = x^2$ around the x-axis (see right graph in the figure below) – namely the volume of the hornlike shape.





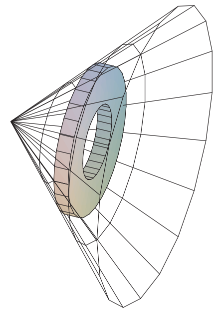
We have already computed the volume of a cone; in this case it is $\pi/3$. At a particular value of x, say x_i, the cross-section of the horn is a circle with radius x_i^2, so the volume of the horn is

$$\int_0^1 \pi(x^2)^2\,dx = \int_0^1 \pi x^4\,dx = \pi\frac{1}{5},$$

so the desired volume is $\pi/3 - \pi/5 = 2\pi/15$.

Method 2: Washer Method

As with the area between curves, there is an alternate approach that computes the desired volume "all at once" by approximating the volume of the actual solid. We can approximate the volume of a slice of the solid with a washer-shaped volume as shown below.



The volume of such a washer is the area of the face times the thickness. The thickness, as usual, is Δx, while the area of the face is the area of the outer circle minus the area of the inner circle, say $\pi R^2 - \pi r^2$. In the present example, at a particular x_i, the radius R is x_i and r is x_i^2. Hence, the whole volume is

$$\int_0^1 \pi x^2 - \pi x^4 \, dx = \pi \left(\frac{x^3}{3} - \frac{x^5}{5} \right) \Bigg|_0^1 = \pi \left(\frac{1}{3} - \frac{1}{5} \right) = \frac{2\pi}{15}.$$

Of course, what we have done here is exactly the same calculation as before. ♣

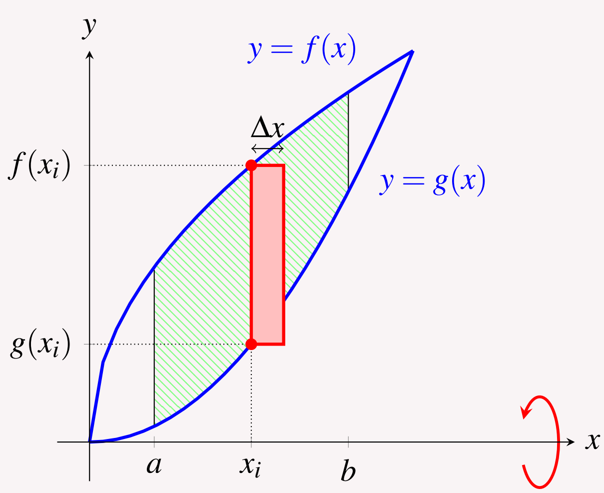
We now formalize the Washer Method employed in the above example.

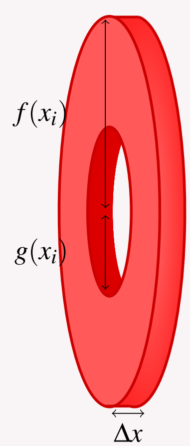
Theorem 3.27: Washer Method: Integration w.r.t. x

Suppose f and g are non-negative and continuous on the interval $[a,b]$ with $f \geq g$ for all x in $[a,b]$. Let R be the area bounded above by f and below by g as well as the lines $x = a$ and $x = b$. Then the volume V formed by rotating R about the x-axis is

$$V = \lim_{\Delta x \to 0} \sum_{i=0}^{n-1} \pi \left([f(x_i)]^2 - [g(x_i)^2] \right) \Delta x = \int_a^b \pi \left([f(x)]^2 - [g(x)^2] \right) dx, \text{ where}$$

x_i is the location of the washer,
$f(x_i)$ is the radius of the outer disk,
$g(x_i)$ is the radius of the inner disk, and
Δx is the thickness of the washer as shown below.





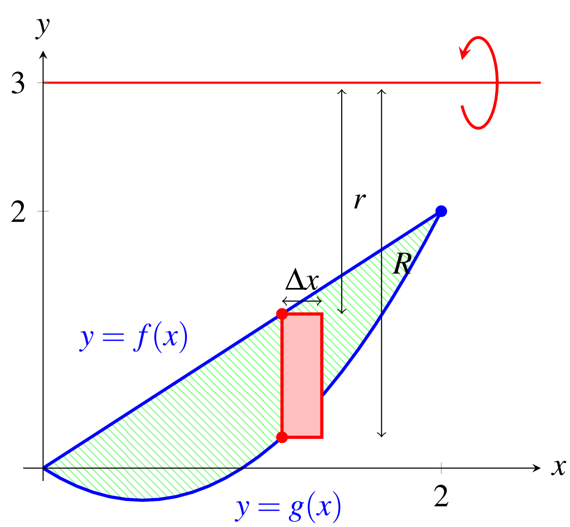
Note: The most common mistake in the Washer Method is to write $(f(x)-g(x))^2$ rather than $f(x)^2 - g(x)^2$.

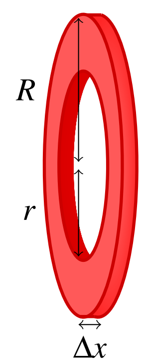
We now present one more example that uses the Washer Method.

Example 3.28: Washer Method

Find the volume of the object generated when the area between $g(x) = x^2 - x$ and $f(x) = x$ is rotated about the line $y = 3$.

Solution. The area between the two curves is graphed below to the left, noting the intersection points $(0,0)$ and $(2,2)$:





From the graph, we see that the inner radius must be $r = 3 - f(x) = 3 - x$, and the outer radius must be $R = 3 - g(x) = 3 - x^2 + x$. Therefore, the volume of the object is

$$\begin{aligned} V &= \int_0^2 \pi \left(\left[3 - x^2 + x\right]^2 - [3-x]^2 \right) dx \\ &= \int_0^2 \pi \left(x^4 - 2x^3 - 6x^2 + 12x\right) dx \\ &= \pi \left[\frac{x^5}{5} - \frac{x^4}{2} - 2x^3 + 6x^2 \right]_0^2 \\ &= \frac{32\pi}{5}. \end{aligned}$$

♣

3.3.4. Disk and Washer Methods: Integration w.r.t. y

We have already seen in Section 3.1 that sometimes a curve is described as a function of y, namely $x = g(y)$, and so the area of the region under the curve g over an interval $[c,d]$ as shown to the left of Figure 3.13 can be rotated about the y-axis to generate a solid of revolution as indicated to the right in Figure 3.13.

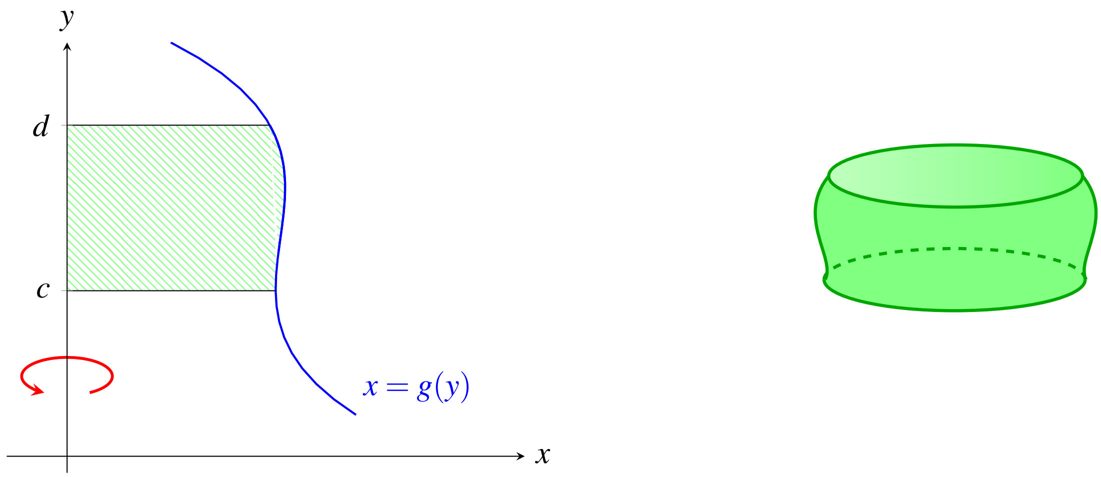


Figure 3.13: Area under a curve $x = g(y)$.

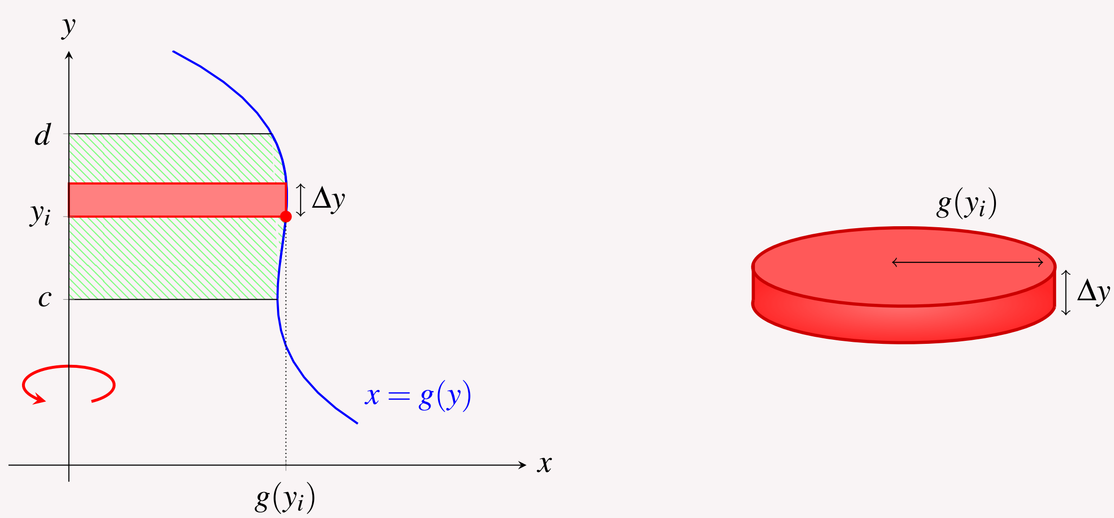
We are readily convinced that the volume of such a solid of revolution can be calculated in a similar manner as those discussed earlier, which is summarized in the following theorem.

Theorem 3.29: Disk Method: Integration w.r.t. y

Suppose g is non-negative and continuous on the interval $[c,d]$. Then the volume V formed by rotating the area under the curve of g about the y-axis is

$$V = \lim_{\Delta y \to 0} \sum_{i=0}^{n-1} \pi \left[g(y_i)\right]^2 \Delta y = \int_a^b \pi \left[g(y)\right]^2 dy, \text{ where}$$

y_i is the location of the disk,
$g(y_i)$ is the radius of the disk, and
Δy is the thickness of the disk as shown below.



Note: Either of the variations 1 and 2 described below will most likely create disks with holes in the solid of revolution, and so we would have to apply the Washer Method.

1. Again, we could rotate the area of any region around an axis of rotation, including the area of a region bounded to the right by a function $x = f(y)$ and to the left by a function $x = g(y)$ on an interval $y \in [c,d]$.

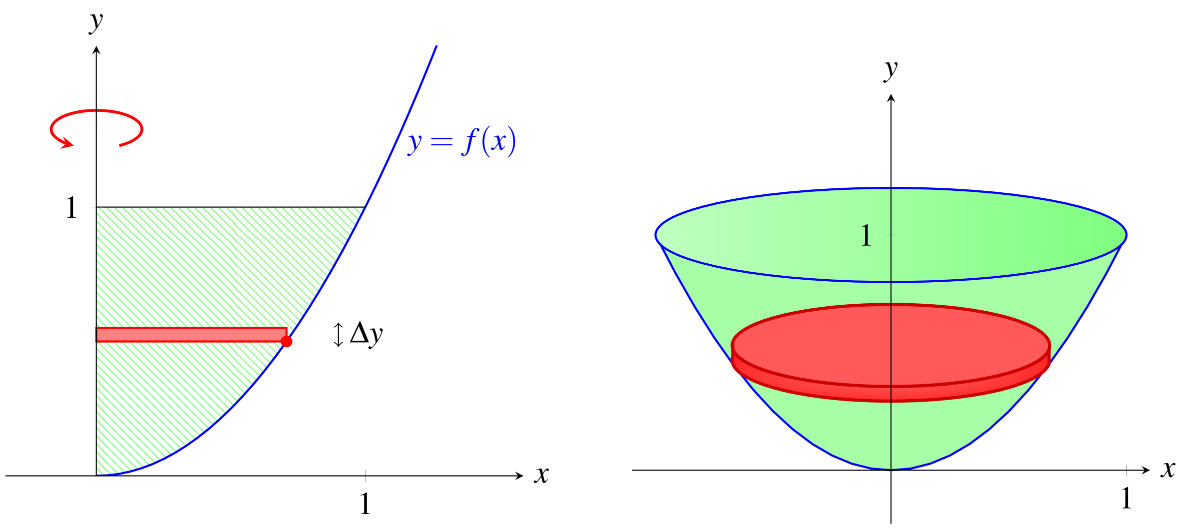
2. The axis of rotation can be any axis parallel to the y-axis for this method to work.

We now provide an example of the Disk Method, where we integrate with respect to y.

Example 3.30: Disk Method

Find the volume of the object generated when the area between the curve $f(x) = x^2$ and the line $y = 1$ in the first quadrant is rotated about the y-axis.

Solution. The area between $y = f(x)$ and $y = 1$ is shown below to the right. We notice that the two curves intersect at $(1,1)$, and that this area is contained between the two curves and the y-axis.



We now solve for x as a function of y:

$$y = x^2 \implies x = \pm\sqrt{y},$$

and since we want the region in the first quadrant, we take $x = \sqrt{y}$. The desired volume is found by integrating

$$\begin{aligned} V &= \pi \int_0^1 (\sqrt{y})^2 \, dy \\ &= \pi \int_0^1 y \, dy \\ &= \frac{\pi y^2}{2}\bigg|_0^1 = \frac{\pi}{2}. \end{aligned}$$

♣

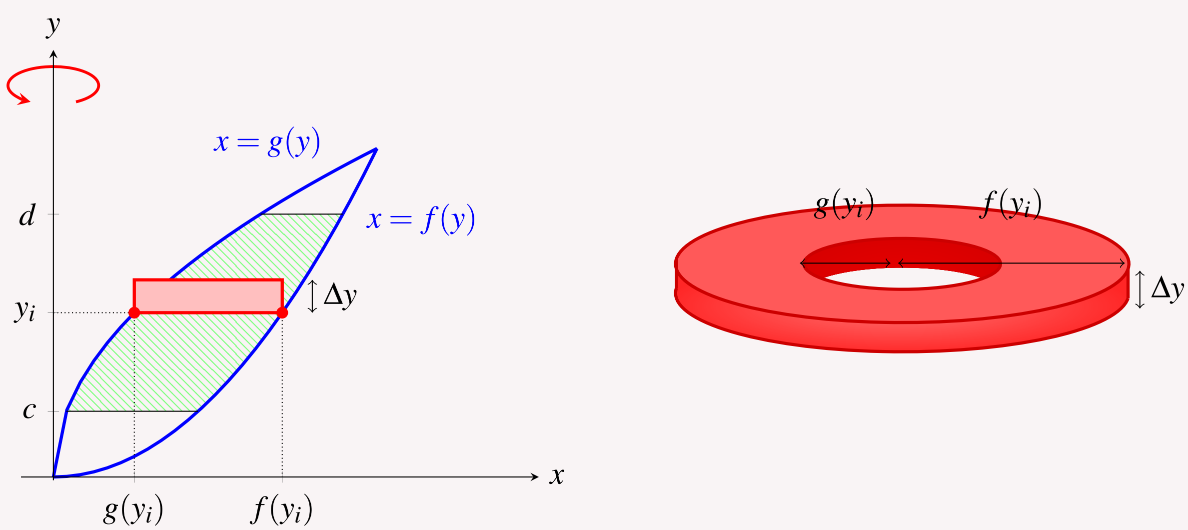
Similar to the Washer Method when integrating with respect to x, we can also define the Washer Method when we integrate with respect to y:

Theorem 3.31: Washer Method: Integration w.r.t. y

Suppose f and g are non-negative and continuous on the interval $[c,d]$ with $f \geq g$ for all y in $[c,d]$. Let R be the area bounded to the right by f and to the left by g as well as the lines $y = c$ and $y = d$. Then the volume V formed by rotating R about the y-axis is

$$V = \lim_{\Delta y \to 0} \sum_{i=0}^{n-1} \pi \left([f(y_i)]^2 - [g(y_i)^2] \right) \Delta y = \int_c^d \pi \left([f(y)]^2 - [g(y)^2] \right) dy, \text{ where}$$

y_i is the location of the washer,
$f(y_i)$ is the radius of the outer disk,
$g(y_i)$ is the radius of the inner disk, and
Δy is the thickness of the washer as shown below.



3.3.5. Summary

There are many different scenarios in which Disk and Washer Methods can be employed, which are not discussed here; however, we provide a general guideline.

Guideline for Disk and Washer Methods

The following steps outline how to employ the Disk or Washer Method.

1. Graph the bounded region.
2. Construct an arbitrary cross-section perpendicular to the axis of rotation.
3. Identify the radius (disk) or radii (washer).
4. Determine the thickness of the disk or washer.
5. Set up the definite integral by making sure you are computing the volume of the constructed cross-section.

Exercises for Section 3.3

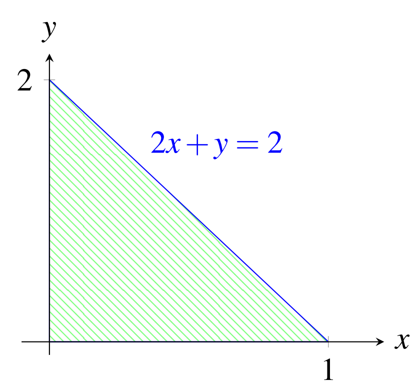
Exercise 3.3.1 *Use the method from Section 3.3.1 to find each volume.*

(a) *The base of a tetrahedron (a triangular pyramid) of height h is an equilateral triangle of side s. Its cross-sections perpendicular to an altitude are equilateral triangles. Express its volume V as an integral, and find a formula for V in terms of h and s. Verify that your answer is $(1/3)$(area of base)(height).*

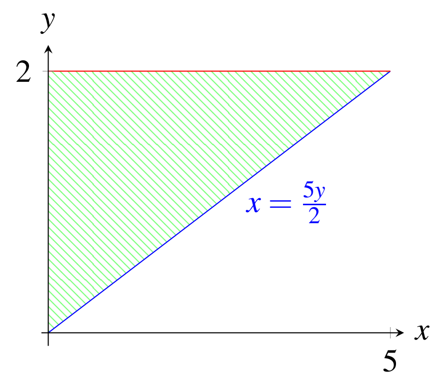
(b) *The base of a solid is the region between $f(x) = \cos x$ and $g(x) = -\cos x$, $-\pi/2 \leq x \leq \pi/2$, and its cross-sections perpendicular to the x-axis are squares. Find the volume of the solid.*

Exercise 3.3.2 *Find the volume of the solid generated by revolving the shaded region about the given axis.*

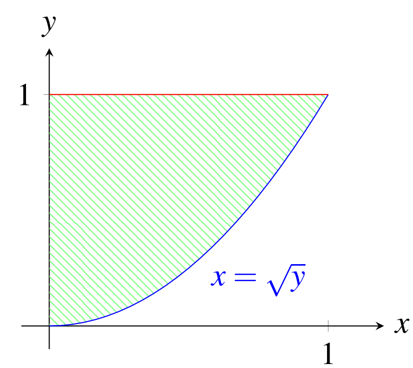
(a) *About the x-axis:*



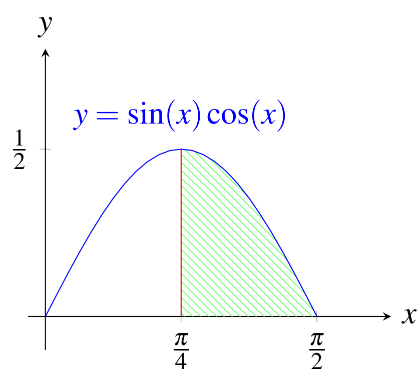
(b) About the y-axis:



(c) About the y-axis:



(d) About the x-axis:



Exercise 3.3.3 *Find the volume of the solid generated by revolving the given bounded region about the x-axis.*

(a) $y = x^2$, $y = 1, x = 2$

(b) $y = x^3$, $y = 0$, $x = 1$

(c) $y = \sqrt{4 - x^2}$, $y = 0$

(d) $y = \sqrt{\sin x}$, $0 \le x \le \pi/2$

(e) $y = e^{-x}$, $y = 0$, $x = 2$

(f) $x + y = 2$, $y = 0$, $x = 0$

(g) $y = x - x^2$, $y = 0$, $x = 0$

(h) $y = \sqrt{\sin x}$, $x = 0$

Exercise 3.3.4 *Find the volume of the solid generated by revolving the given bounded region about the y-axis.*

(a) $x = y^2$, $x = 0$, $y = -1$, $y = 1$

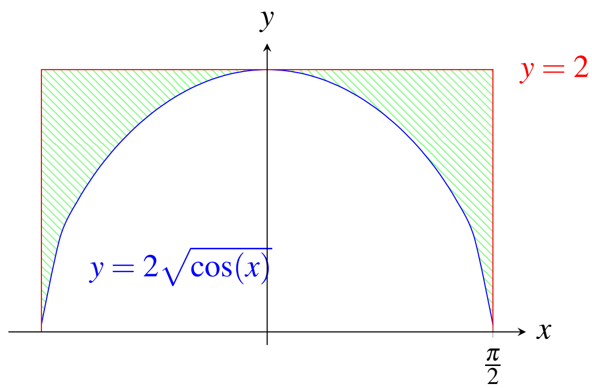
(b) $x = y^{3/2}$, $x = 0$, $y = 4$

(c) $x = \sqrt{\sin(2y)}$, $0 \leq y \leq \pi/2$, $x = 0$

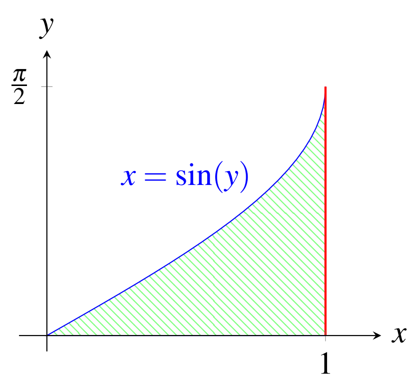
(d) $x = \sqrt{\cos(2y)}$, $0 \leq y \leq \pi/2$, $x = 0$

Exercise 3.3.5 *Find the volume of the solid generated by revolving the shaded region about the given axis.*

(a) The x-axis:



(b) The y-axis:



Exercise 3.3.6 *Find the volume of the solid generated by revolving the given bounded region about the x-axis.*

(a) $y = x, y = 2, x = 0$

(b) $y = 3\sqrt{x}, y = 3, x = 0$

(c) $y = x^2 + 1, y + x = 3$

(d) $y = 9 - x^2, y = 3 - x$

Exercise 3.3.7 *The equation $x^2/9 + y^2/4 = 1$ describes an ellipse. Find the volume of the solid obtained by rotating the ellipse around the x-axis and also around the y-axis. These solids are called* **ellipsoids***; one is vaguely rugby-ball shaped, one is sort of flying-saucer shaped, or perhaps squished-beach-ball-shaped.*

Figure 3.14: Ellipsoids.

Exercise 3.3.8 *Use integration to compute the volume of a sphere of radius r. You should of course get the well-known formula* $4\pi r^3/3$.

Exercise 3.3.9 *A hemispheric bowl of radius r contains water to a depth h. Find the volume of water in the bowl.*

3.4 Volume of Revolution: Shell Method

In the previous section, we calculated the volume of a solid of revolution over a closed interval $[a,b]$ by adding up the cross-sectional areas, which we obtained by slicing through the solid with planes perpendicular to the axis of rotation over $[a,b]$. There is another method, which instead of creating *disk-slices* will create *cylindrical shell-slices*. These cylindrical shell-slices are created by cutting through the solid with cylinders that wrap symmetrically around the axis of rotation as shown in Figure 3.15. This is similar to stacking paper towel rolls of increasing radii inside of each other.

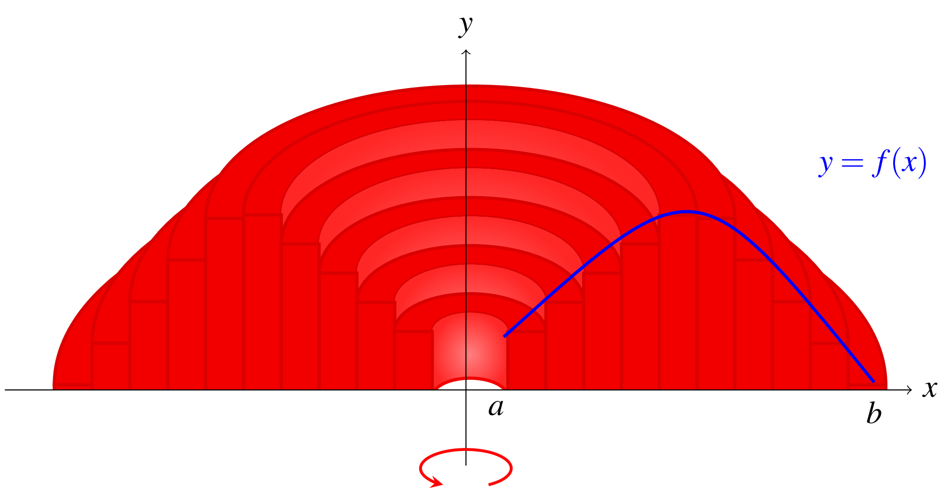
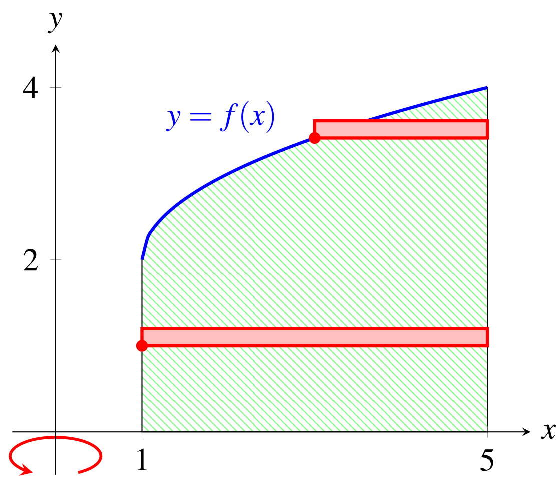


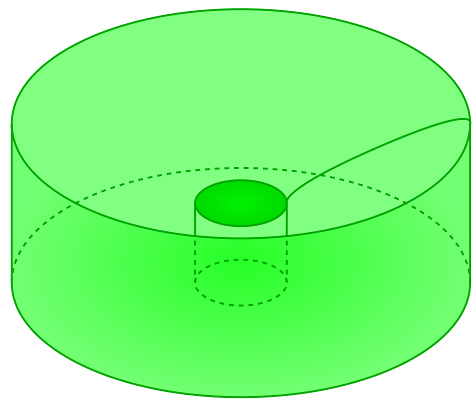
Figure 3.15: Cylindrical Shells

Just like we were able to add up disks, we can also add up **cylindrical shells**, and therefore this method of integration for computing the volume of a solid of revolution is referred to as the **Shell Method**. We begin by investigating such shells when we rotate the area of a bounded region around the y-axis.

3.4.1. Shell Method: Integration w.r.t. x

Suppose the region bounded by $f(x) = \sqrt{x-1}+2$ with $x \in [1,5]$ is rotated around the y-axis as shown below to the right. It is possible, but inconvenient, to compute the volume of the resulting solid by the Washer Method we have used so far. The problem is that there are two "kinds" of typical washers: Those that go from the curve f to the line $x=5$ and those that sit between the lines $x=1$ and $x=5$ as shown below to the left.

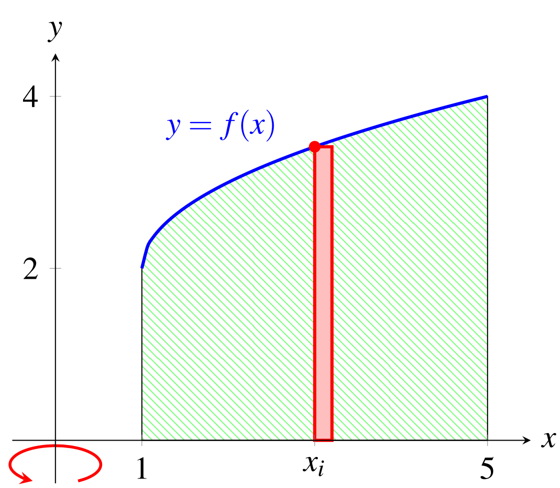


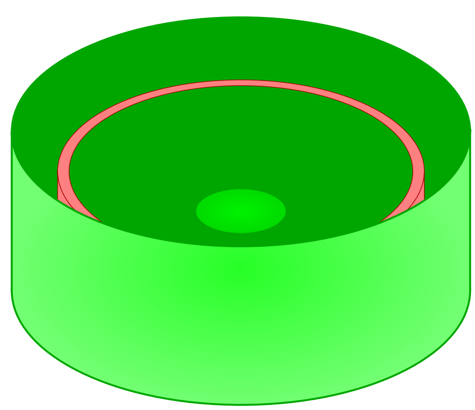


To compute the volume using this approach, we need to break the problem into two parts and compute two integrals:

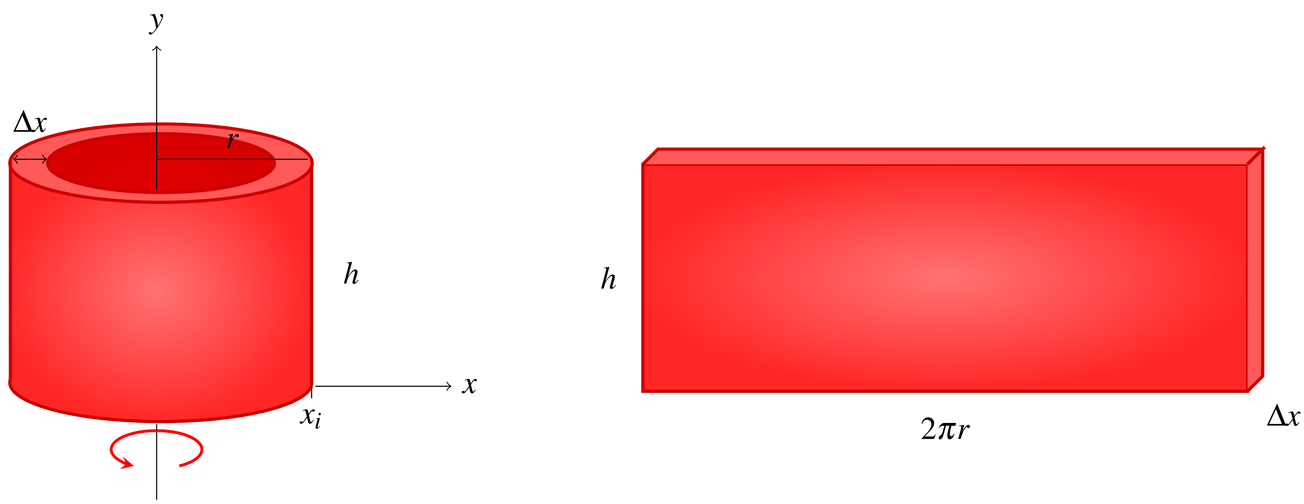
$$\begin{aligned} V &= \int_0^2 \pi(5^2-1^2)\,dy + \int_2^4 \pi\left[5^2 - \left((y-2)^2+1\right)^2\right]dy \\ &= \int_0^2 24\pi\,dy + \int_2^4 \pi\left(-y^4+8y^3-26y^2+40y\right)dy \\ &= 24\pi y\Big|_0^2 + \pi\left[-\frac{y^5}{5}+2y^4-\frac{26y^3}{3}+20y^2\right]_2^4 = 84\frac{4}{15}\pi \end{aligned}$$

If instead we slice through the solid of revolution parallel to the axis of rotation using cylindrical shells with increasing radii over the interval $[1,5]$ along the x-axis, then any of the cylindrical shells has height f as shown below.





Note that "washers" are related to the area of a circle, πr^2, whereas "shells" are related to the surface area of an open cylinder, $2\pi rh$. If we add up the volume of such thin shells we will get an approximation to the true volume. What is the volume of such a shell? Consider the shell at x_i. Imagine that we cut the shell vertically in one place and "unroll" the surface of the cylindrical shell into a thin, flat sheet, as shown below.



This sheet will be almost a rectangular prism that is Δx thick, h high, and $2\pi r$ wide (the circumference of the cylindrical shell). In terms of x_i, notice that h is the height of f, and that the radius r is precisely x_i. The volume of one cylindrical shell will then be approximately the volume of a rectangular prism with these dimensions: $2\pi x_i f(x_i)\Delta x$. If we add these up and take the limit as usual, we get the integral

$$V = \int_1^5 2\pi x f(x)\,dx = \int_1^5 2\pi x\left(\sqrt{x-1}+2\right)\,dx.$$

Now use substitution to evaluate the integral:

$$u = x-1,\ x = u+1,\ du = dx,\ u(1) = 0,\ u(5) = 4.$$

Therefore,

$$\begin{aligned} V &= \int_1^5 2\pi x\left(\sqrt{x-1}+2\right)\,dx \\ &= 2\pi\int_0^4 (u+1)\left(\sqrt{u}+2\right)\,du \\ &= 2\pi\int_0^4 \left(u^{3/2}+2u+u^{1/2}+2\right)\,du \\ &= 2\pi\left[\frac{2u^{5/2}}{5}+u^2-\frac{2u^{3/2}}{3}+2u\right]_0^4 \\ &= 84\frac{4}{15}\pi. \end{aligned}$$

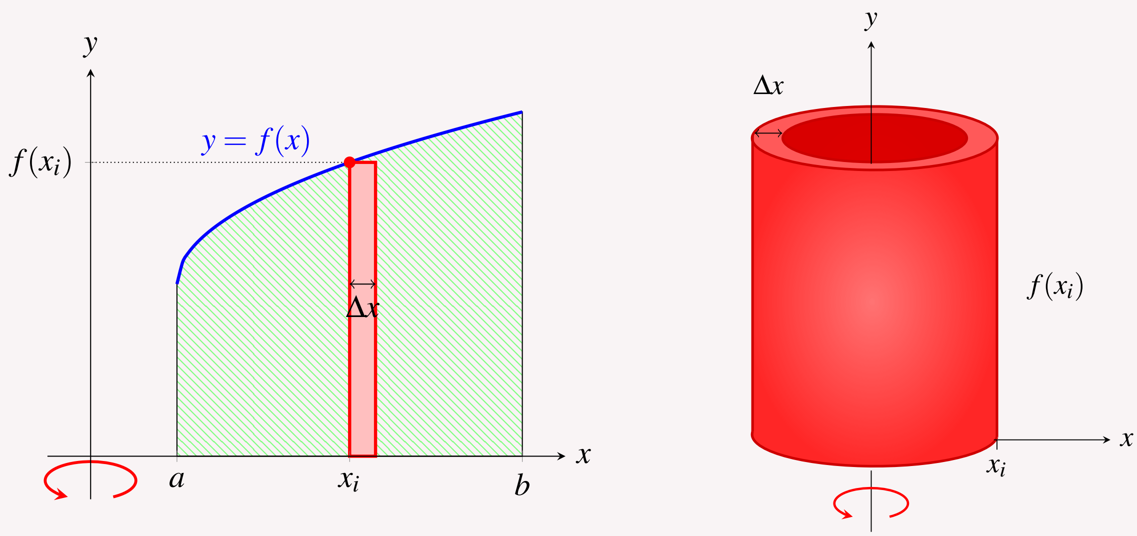
This accomplishes the task with only one integral. One may argue that in the end, we had to do the same amount of work. However, there are often situations when one of the two methods, washer or shell, is easier than the other. Therefore, it is worthwhile investigating this technique when computing volumes of solids of revolution. We capture our results in the following theorem.

Theorem 3.32: Shell Method: Area under Curve – Integration w.r.t. x

Suppose f is non-negative and continuous on the interval $[a,b]$. Then the volume V formed by rotating the area under the curve of f about the y-axis is

$$V = \lim_{\Delta x \to 0} \sum_{i=0}^{n-1} 2\pi x_i f(x_i)\Delta x = \int_a^b 2\pi x f(x)\, dx, \text{ where}$$

x_i is the location of the cylindrical shell and its radius,
$f(x_i)$ is the height of the cylindrical shell, and
Δx is the thickness of the cylindrical shell as shown below.

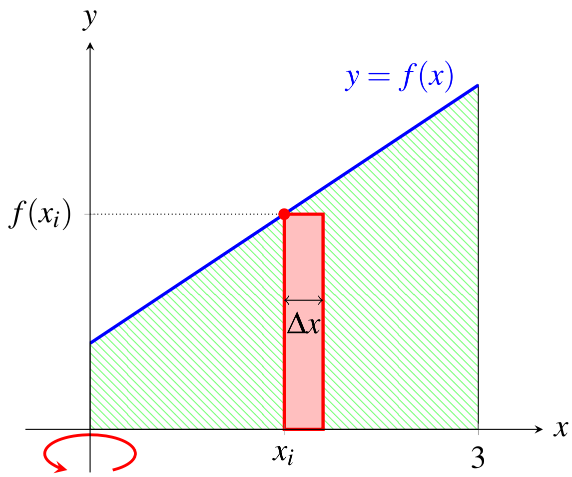


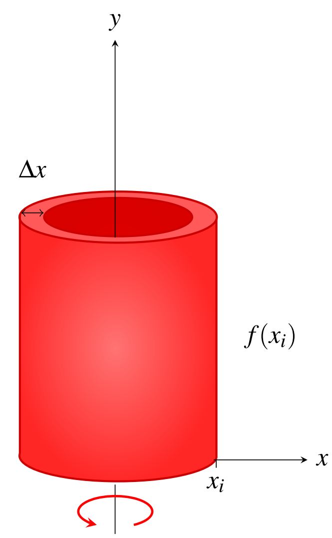
We now provide one example of the Shell Method.

Example 3.33: Shell Method – Integration w.r.t. x

Suppose the area below the curve $f(x) = x+1$ for all x in $[0,3]$ is rotated about the y-axis. Find the volume using the Shell Method.

Solution. We begin by graphing f on the interval $[0,3]$ and identify an arbitrary cylindrical shell as shown below.





Then the volume is given by

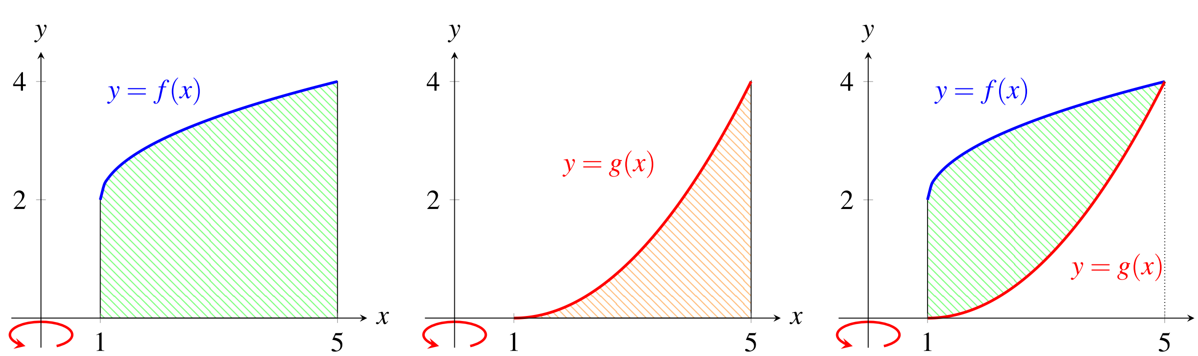
$$V = \int_0^3 2\pi x f(x)\,dx = \int_0^3 2\pi x(x+1)\,dx$$
$$= 2\pi \int_0^3 (x^2+x)\,dx = 2\pi \left[\frac{x^3}{3} + \frac{x^2}{2}\right]_0^3 = 27\pi.$$

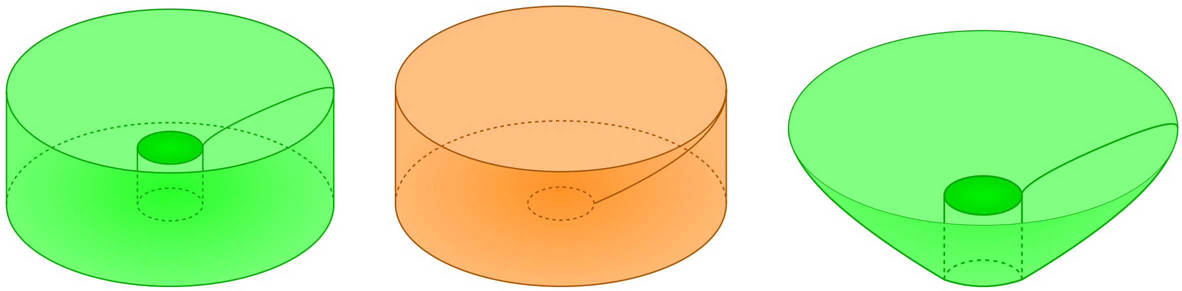
♣

Note:

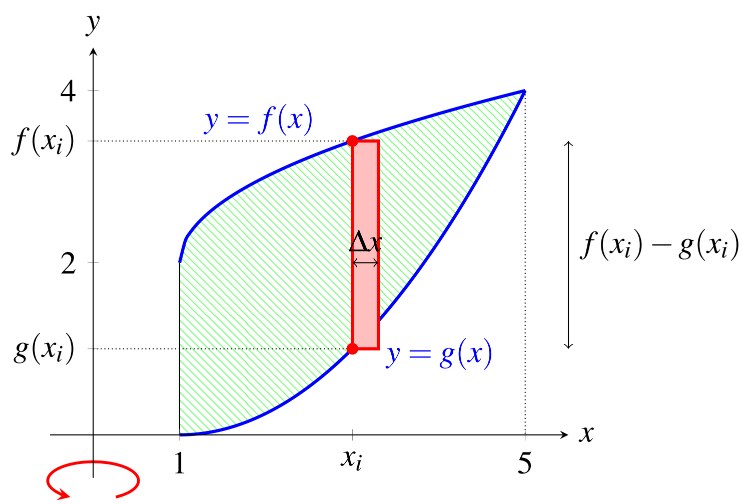
1. Once again, we can rotate the area of any region around an axis of rotation, including the area of a region bounded above by a function $y = f(x)$ and below by a function $y = g(x)$ on an interval $x \in [a,b]$ as we will investigate next.

2. The axis of rotation can be any axis parallel to the y-axis for this method to work. Consider again the region bounded by the function $f(x) = \sqrt{x-1}+2$ with $x \in [1,5]$, but let's also include a second bound, namely the function $g(x) = \frac{1}{4}(x-1)^2$. Algebra shows us that $f > g$ on $x \in [1,5]$. We want to compute the volume of the solid obtained by rotating the bounded region about the y-axis. We begin by graphing the region and remind ourselves that the given region is the difference of two regions as shown below:





The above figure shows us that the volume of the solid of revolution obtained from rotating the given region is simply the difference between the volumes of the two solids of revolution that are created from rotating the area under the curve of f and the area under the curve of g respectively. An arbitrary cylindrical shell must therefore have height $f(x_i) - g(x_i)$ as shown below.



Therefore, the volume is given by

$$V = \int_1^5 2\pi x\,[f(x) - g(x)]\,dx = \int_1^5 2\pi x\left[\sqrt{x-1} + 2 - \frac{(x-1)^2}{4}\right] dx = \frac{208\pi}{5}$$

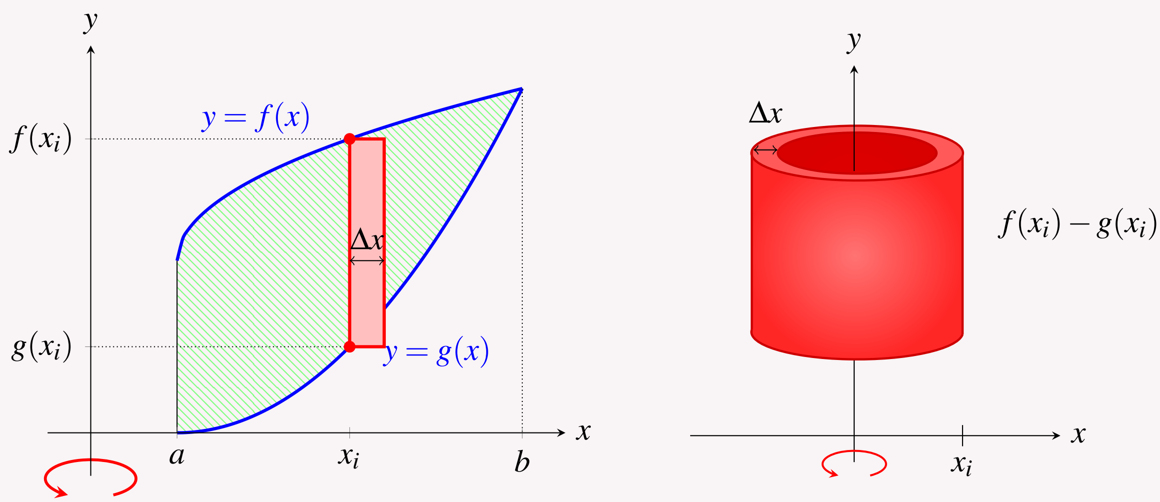
We capture our result in the following theorem.

Theorem 3.34: Shell Method: Area between Curves – Integration w.r.t. x

Suppose f and g are non-negative and continuous on the interval $[a,b]$ with $f \geq g$ for all x in $[a,b]$. Let R be the area bounded above by f and below by g as well as the lines $x=a$ and $x=b$. Then the volume V formed by rotating R about the y-axis is

$$V = \lim_{\Delta x \to 0} \sum_{i=0}^{n-1} 2\pi x_i [f(x_i) - g(x_i)]\, \Delta x = \int_a^b 2\pi x [f(x) - g(x)]\, dx, \text{ where}$$

x_i is the location of the cylindrical shell and its radius,
$f(x_i) - g(x_i)$ is the height of the cylindrical shell, and
Δx is the thickness of the cylindrical shell as shown below.



We now provide one more example of such a region bounded below and above by two functions f and g respectively.

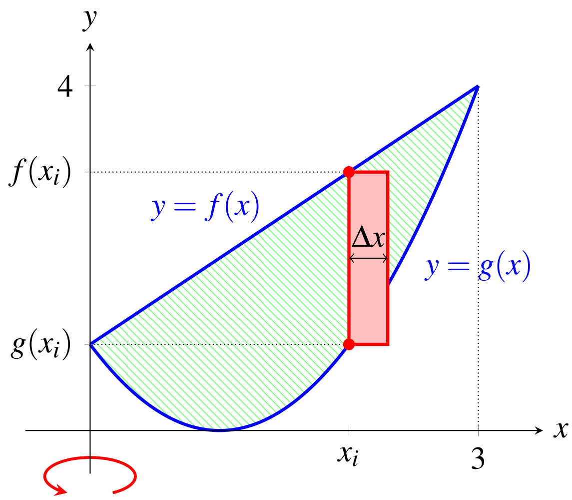
Example 3.35: Shell Method – Integration w.r.t x

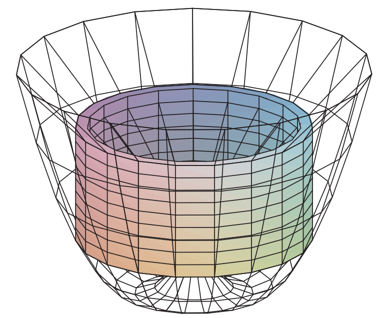
Suppose the region between $f(x) = x+1$ and $g(x) = (x-1)^2$ is rotated around the y-axis. Find the volume using the Shell Method.

Solution. We begin by finding the intersection points of the two functions f and g:

$$\begin{aligned} f(x) &= g(x) \\ x+1 &= (x-1)^2 \\ 0 &= x^2 - 3x = x(x-3) \end{aligned}$$

and so $x=0$ and $x=3$ are the intersection points. We now use this information to graph f and g on the interval $[0,3]$ and identify an arbitrary cylindrical shell as shown below.





Then the volume is computed as follows:

$$\begin{aligned} V &= \int_0^3 2\pi x\,[f(x) - g(x)]\,dx \\ &= \int_0^3 2\pi x\left[(x+1) - (x-1)^2\right]\,dx \\ &= 2\pi \int_0^3 (3x^2 - x^3)\,dx \\ &= 2\pi \left[x^3 - \frac{x^4}{4}\right]_0^3 \\ &= \frac{27}{2}\pi \end{aligned}$$

♣

3.4.2. Shell Method: Integration w.r.t. y

So far, we have discussed three main manners of generating a solid of revolution and how to compute its volume, which are listed below. Remember that the Washer Method is replaced by the Disk Method when the lower or left curve is described by the x-axis or the y-axis respectively.

- Rotating an area that is bounded above and below by functions of x as well as lines $x = a$ and $x = b$ around the x-axis, and then using the Washer Method for volume-computation.
- Rotating an area that is bounded right and left by functions of y as well as lines $y = c$ and $y = d$ around the y-axis, and then using the Washer Method for volume-computation.
- Rotating an area that is bounded above and below by functions of x as well as lines $x = a$ and $x = b$ around the y-axis, and then using the Shell Method for volume-computation.

There is only one case left:

- Rotating an area that is bounded right and left by functions of y as well as lines $y = c$ and $y = d$ around the y-axis, and then using the Shell Method for volume-computation.

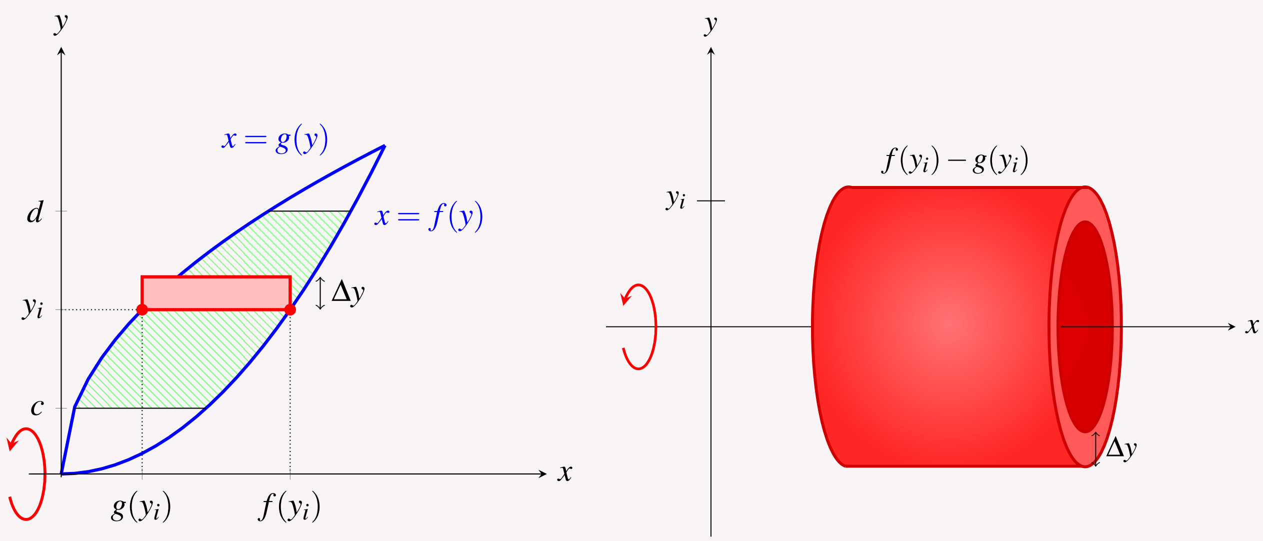
We are readily convinced that the volume of such a solid of revolution can be calculated using a Shell Method similar in manner as the one discussed earlier, which is summarized in the following theorem.

Theorem 3.36: Shell Method: Area between Curves – Integration w.r.t. y

Suppose f and g are non-negative and continuous on the interval $[c,d]$ with $f \geq g$ for all y in $[c,d]$. Let R be the area bounded right by f and left by g as well as the lines $y = c$ and $y = d$. Then the volume V formed by rotating R about the x-axis is

$$V = \lim_{\Delta y \to 0} \sum_{i=0}^{n-1} 2\pi y_i \left[f(y_i) - g(y_i)\right] \Delta y = \int_c^d 2\pi y \left[f(y) - g(y)\right] dy, \text{ where}$$

y_i is the location of the cylindrical shell and its radius,
$f(y_i) - g(y_i)$ is the height of the cylindrical shell, and
Δy is the thickness of the cylindrical shell as shown below.



We provide one example to exemplify this method as well.

Example 3.37: Comparing Methods

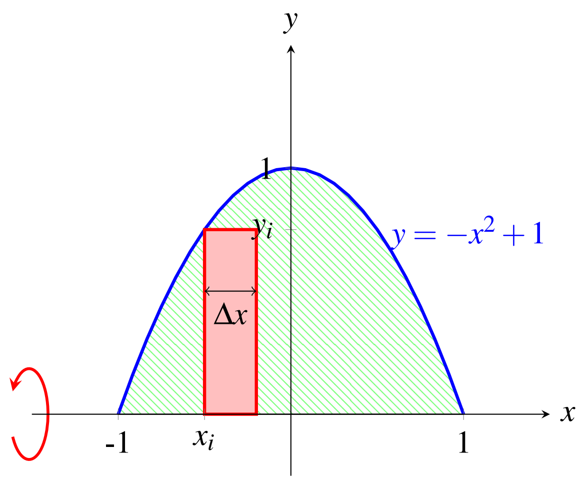
Suppose the area under $y = -x^2 + 1$ is rotated around the x-axis. Find the volume of the solid of rotation using

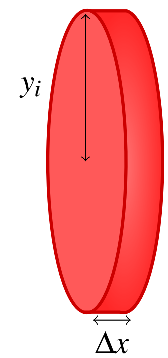
(a) the Disk Method; and

(b) the Shell Method.

Solution.

(a) We begin by graphing the area and indicate one arbitrary disk:

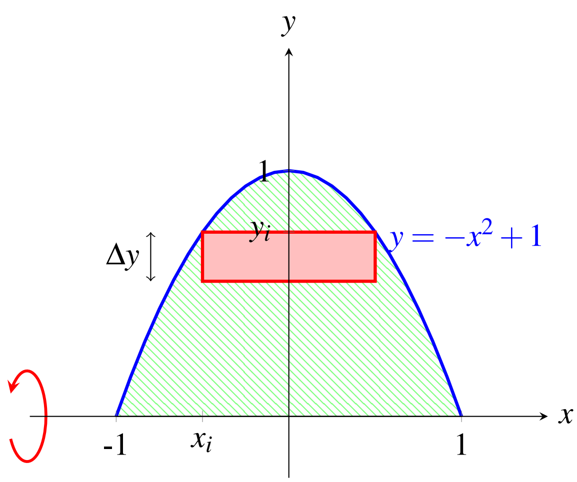


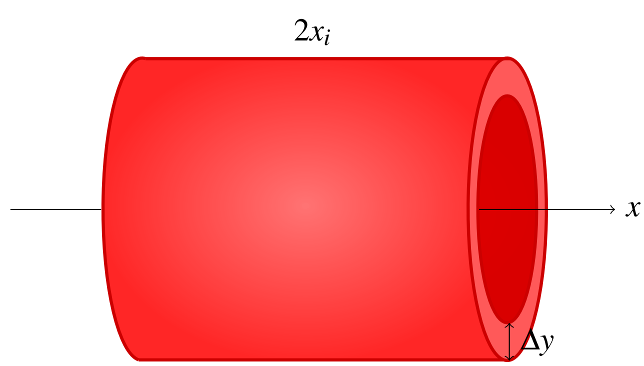


Therefore, the volume is given by

$$\begin{aligned} V &= \int_{-1}^{1} \pi y^2 \, dx = \int_{-1}^{1} \pi \left(1 - x^2\right)^2 dx \\ &= 2\pi \int_0^1 \left(x^4 - 2x^2 + 1\right) dx = 2\pi \left[\frac{x^5}{5} - \frac{2x^3}{3} + x\right]_0^1 = \frac{16}{15}\pi \end{aligned}$$

(b) We again begin by graphing the area and indicate one arbitrary cylindrical shell:





First, we have to express x in terms of y:

$$y = 1 - x^2 \implies x = \pm\sqrt{1 - y}$$

Therefore, the volume is computed with

$$V = \int_0^1 2\pi y(2x) \, dy = 4\pi \int_0^1 y\sqrt{1-y} \, dy.$$

Now we use substitution

$$u = 1 - y,\ y = 1 - u,\ du = -dy,\ u(0) = 1,\ u(1) = 0$$

and get

$$V = 4\pi \int_0^1 y\sqrt{1-y}\,dy = 4\pi \int_1^0 (1-u)\sqrt{u}\,(-du) = 4\pi \int_0^1 \left(u^{1/2} - u^{3/2}\right) du$$

$$= 4\pi \left[\frac{2u^{3/2}}{3} - \frac{2u^{5/2}}{5}\right]_0^1 = \frac{16}{15}\pi$$

♣

3.4.3. Summary

There are many different scenarios in which the Shell Method can be employed, which are not discussed here; however, we provide a general guideline.

Guideline for Shell Method

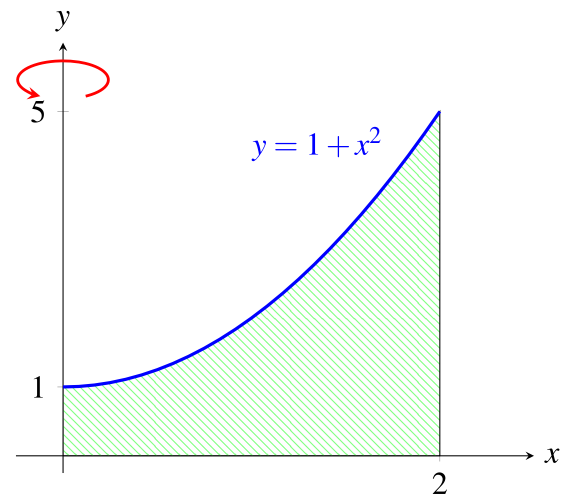
The following steps outline how to employ the Shell Method.

1. Graph the bounded region.
2. Construct an arbitrary cylindrical shell parallel to the axis of rotation.
3. Identify the radius and height of the cylindrical shell.
4. Determine the thickness of the cylindrical shell.
5. Set up the definite integral by making sure you are computing the volume of the constructed cylindrical shell.

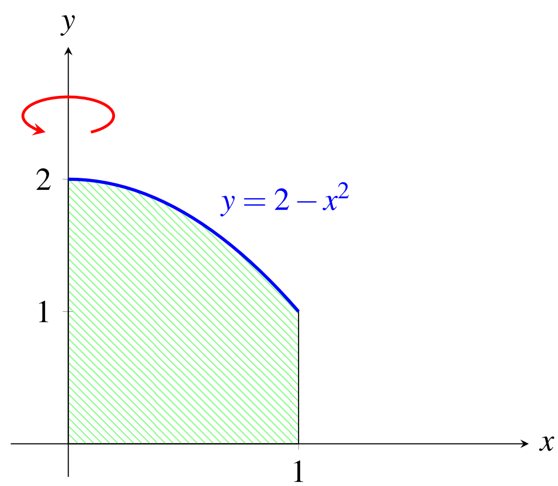
Exercises for Section 3.4

Exercise 3.4.1 *Use Shell Method to find the volumes of the solids generated by revolving the shaded region about the indicated axis.*

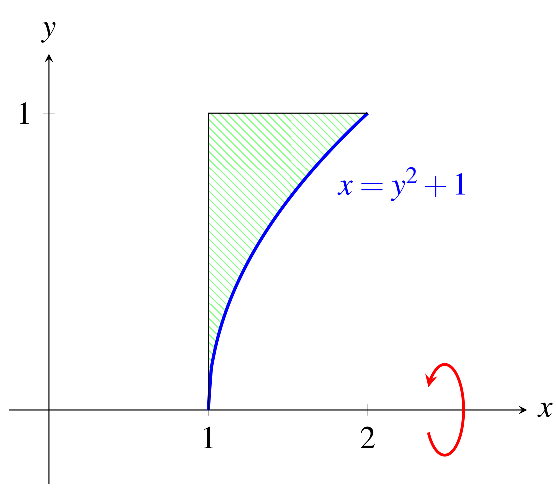
(a) The y-axis:



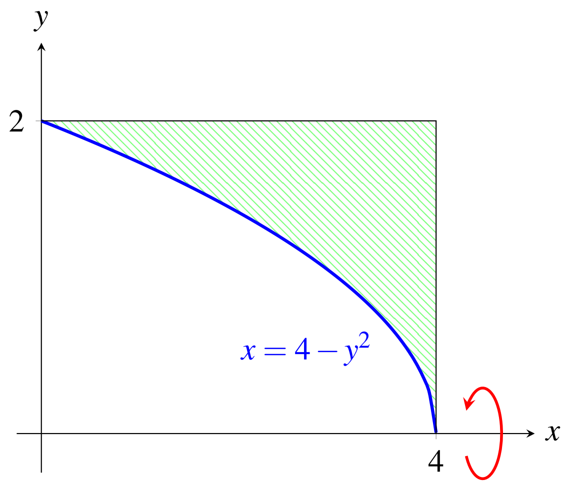
(b) *The y-axis:*



(c) *The x-axis:*



(d) *The x-axis:*



Exercise 3.4.2 *Use Shell Method to find the volumes of the solids generated by revolving the given bounded regions about the y-axis.*

(a) $y = 2x,\ y = -x,\ x = 1$

(b) $y = x,\ y = 2x,\ x = 2$

(c) $y = x^2,\ y = x + 2,\ x \geq 0$

(d) $y = 1 - x^2,\ x^2,\ x \geq 0$

(e) $y = \sqrt{x},\ x - 1,\ x = 0$

(f) $y = \frac{1}{x},\ x = 2,\ x = 3,\ y = 0$

Exercise 3.4.3 *Use Shell Method to find the volumes of the solids generated by revolving the given bounded regions about the x-axis.*

(a) $x = 2\sqrt{y},\ x = -y,\ y = 4$

(b) $x = y^2,\ x = y,\ y \geq 0$

(c) $x = y - y^2/4,\ x = 0$

(d) $x = y - y^2/4,\ x = y/2$

(e) $y = |x|,\ y = 2$

(f) $y = x + 1,\ y = 2x$

Exercise 3.4.4 *Let R be the region of the x-y-plane bounded below by the curve $y = \sqrt{x}$, and above by the line $y = 3$. Find the volume of the solid obtained by rotating R around*

(a) *the x-axis;*

(b) *the line $y = 3$;*

(c) *the y-axis, and*

(d) *the line $x = 9$.*

Exercise 3.4.5 *Let S be the region of the x-y-plane bounded above by the curve $x^3y = 64$, below by the line $y = 1$, on the left by the line $x = 2$, and on the right by the line $x = 4$. Find the volume of the solid obtained by rotating S around* *(a) the x-axis;*

(b) the line $y = 1$*;*

(c) the y-axis, and

(d) the line $x = 2$.

4. Multiple Integration

4.1 Functions of Several Variables

Recall that a function $f:\mathbb{R}\to\mathbb{R}$ maps a single real value x to a single real value y. Such a function is referred to as a **single-variable** function and can be readily visualized in a two-dimensional coordinate system: above (or below) each point x on the x-axis we graph the point y, where of course $y=f(x)$. By now, you have seen the graphs of many such functions. We now extend this visualization process to **multi-variable** functions, also referred to as functions of several variables.

In single-variable calculus we were concerned with functions that map the real numbers $\mathbb{R}$ to $\mathbb{R}$, sometimes called "real functions of one variable", meaning the "input" is a single real number and the "output" is likewise a single real number. Now we turn to functions of several variables, where several input variables are mapped to one value: functions $f\colon\mathbb{R}^n\to\mathbb{R}$. We will deal primarily with $n=2$ and to a lesser extent $n=3$; in fact many of the techniques we discuss can be applied to larger values of n as well.

A function $f\colon\mathbb{R}^2\to\mathbb{R}$ maps a pair of values (x,y) to a single real number. The three-dimensional coordinate system is a convenient way to visualize such functions: above (or below) each point (x,y) in the x-y-plane we graph the point (x,y,z), where then $z=f(x,y)$. In other words, in interpreting the graph of a function $f(x,y)$, one often thinks of the value $z=f(x,y)$ of the function at the point (x,y) as the *height* of the point (x,y,z) on the graph of f. If $f(x,y)>0$, then the point (x,y,z) is $f(x,y)$ units above the x-y-plane; if $f(x,y)<0$, then the point (x,y,z) is $|f(x,y)|$ units below the x-y-plane.

In general, it is quite difficult to draw the graph of a function of two variables. But techniques have been developed that enable us to generate such graphs with minimum effort, using a computer. Still, it is valuable to be able to visualize relatively simple surfaces without such aids. It is often a good idea to examine the function on restricted subsets of the plane, especially lines. It can also be useful to identify those points (x,y) that share a common z-value.

Before we introduce some special and some general three-dimensional graphs and their equations, let us first consider the domain of functions in two variables.

The variables x and y are called *independent* variables, and the variable z, which is dependent on the values of x and y, is referred to as a *dependent* variable. As indicated already, the number $z=f(x,y)$ is called the *value* of f at the point (x,y). Unless specified otherwise, the domain of the function f will be taken to be the largest possible set for which the rule defining f is meaningful. We will demonstrate this with several examples.

Example 4.1: Domain of Two-Variable Functions

Find the domain of each of the following functions.

(a) $f(x,y) = \sqrt{1-x^2-y^2}$.

(b) $g(x,y) = \dfrac{1}{\sqrt{1-x^2-y^2}}$.

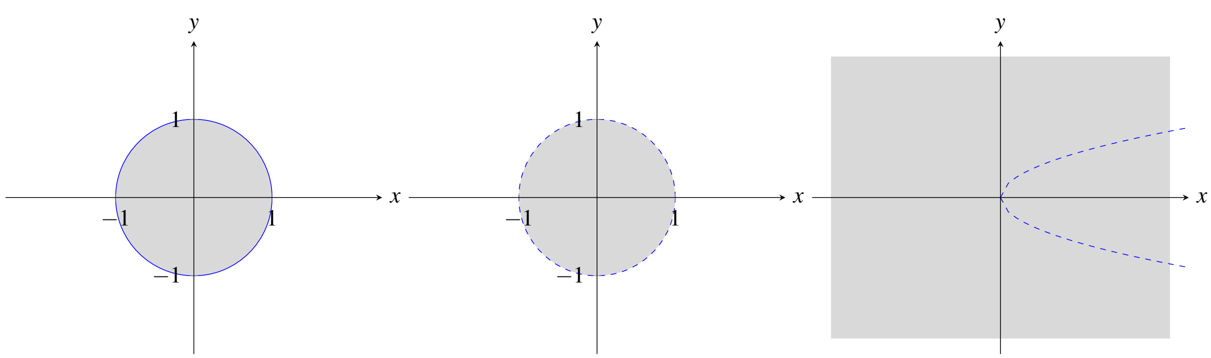
(c) $h(x,y) = \dfrac{1}{y-x^2}$.

Solution.

(a) The domain of f is the set of points (x,y) such that $1-x^2-y^2 \geq 0$. We recognize $x^2+y^2=1$ as the equation of a circle of radius 1 centred at the origin, and so the domain of f consists of all points which lie *on* or *inside* this circle (see below).

(b) We now require that $1-x^2-y^2 > 0$. The domain of g therefore contains all points (x,y) such that $x^2+y^2 < 1$; that is, all points which lie strictly *inside* the unit circle (see below).

(c) We see that $h(x,y)$ is undefined for $x = y^2$. The domain of h therefore consists of all points in the x-y-plane except those which satisfy $y = \pm\sqrt{x}$ (see below).



(a) *Domain of f shown in grey.* (b) *Domain of g shown in grey.* (c) *Domain of h shown in grey.*

♣

Example 4.2: Domain of Two-Variable Functions

A manufacturer produces a model X and a model Y, and determines that the unit prices of these two products are related. Let q_x be the weekly quantity demanded of model X, and let q_y be the weekly quantity demanded of model Y. The unit price of model X is found to be

$$p_x = 500 - q_x - \frac{1}{3}q_y,$$

and the unit price of model Y is found to be

$$p_y = 200 - \frac{1}{3}q_x - \frac{1}{5}q_y.$$

(a) Determine the revenue function $R(q_x, q_y)$.

(b) Sketch the domain of R.

Solution.

(a) Selling q_x units of model X yields a revenue of $q_x p_x$ dollars per week, and selling q_y units of model Y yields a revenue of $q_y p_y$ dollars per week. We therefore construct the revenue function R as

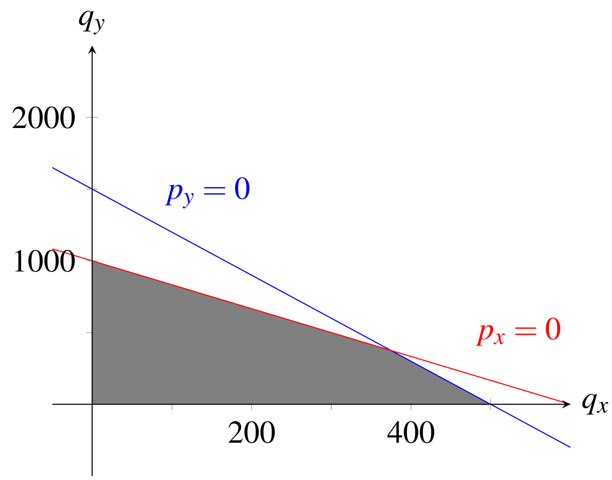
$$\begin{aligned} R(q_x, q_y) &= q_x p_x + q_y p_y \\ &= q_x\left(500 - q_x - \frac{1}{3}q_y\right) + q_y\left(200 - \frac{1}{3}q_x - \frac{1}{5}q_y\right) \\ &= \frac{1}{15}\left(-15q_x^2 - 10q_xq_y + 7500q_x - 3q_y^2 + 3000q_y\right), \end{aligned}$$

dollars per week.

(b) The domain of R is all points (q_x, q_y) in the plane such that q_x,q_y,p_x and $p_y \geq 0$. That is, where

$$\begin{aligned} q_x &\geq 0, \\ q_y &\geq 0, \\ 500 - q_x - \tfrac{1}{3}q_y \geq 0 &\iff q_y \leq 1500 - 3q_x \\ 200 - \tfrac{1}{3}q_x - \tfrac{1}{5}q_y \geq 0 &\iff q_y \leq 1000 - \tfrac{5}{3}q_x. \end{aligned}$$

We find the domain by graphing the two lines given by $p_x = 0$ and $p_y = 0$ in the first quadrant of the q_x-q_y-plane. The domain of R is the region which is bounded by the lines $p_x = 0$, $p_y = 0$, $q_x = 0$ and $q_y = 0$ for which the desired inequality holds. The final solution is given by the shaded region (including the boundary) in the figure below.



♣

Let us now consider some functions that we can describe and graph with the knowledge we have attained so far.

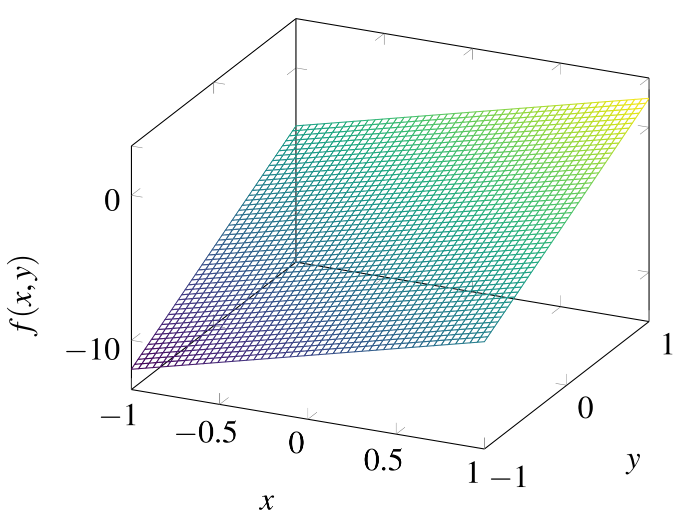
Example 4.3: Plane

Describe and graph the function $f(x,y) = 3x+4y-5$.

Solution.

We first write this as $z = 3x+4y-5$ and then $3x+4y-z = 5$, which is the equation of a plane. However, we can also determine that this equation represents a plane from the following analysis: If we hold x constant, then the equation simplifies to that of a straight line in the y-z-plane. Generally speaking, if we choose any point on the graph of f, for example from $f(0,0) = -5$, this function grows linearly in every direction in the same way that a linear function behaves as can be seen below.

To graph the plane, we alternately let x, then y, then z be equal to zero, which leads to a linear equation in the y-z-, x-z-, and x-y-planes respectively. These lines represent the intersection of the functions' graph with each of the coordinate planes in the three-dimensional coordinate system and form part of the plane surface as shown below.

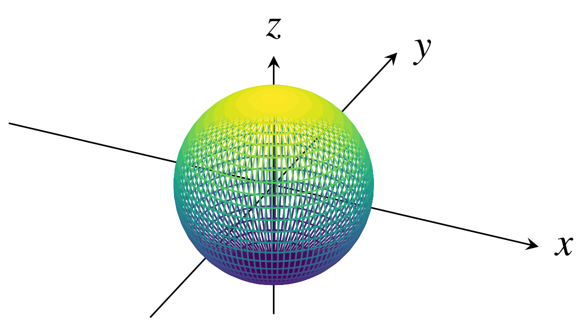


♣

Example 4.4: Sphere

Describe and graph the equation $x^2+y^2+z^2=4$.

Solution. The equation $x^2+y^2+z^2=4$ represents a sphere of radius 2 and centre $(0,0,0)$ as shown below. We cannot write this in the form $z=f(x,y)$, since for each x and y in the disk $x^2+y^2<4$ there are two corresponding points on the sphere, namely one above and one below this point (x,y). As with the equation of a circle, we can resolve this equation into two functions, $f_1(x,y)=\sqrt{4-x^2-y^2}$ and $f_2(x,y)=-\sqrt{4-x^2-y^2}$, representing the upper and lower hemispheres, respectively. Each of these is an example of a function with a restricted domain: only certain values of x and y make sense (namely, those for which $x^2+y^2\le 4$) and the graphs of these functions are limited to a small region of the plane.

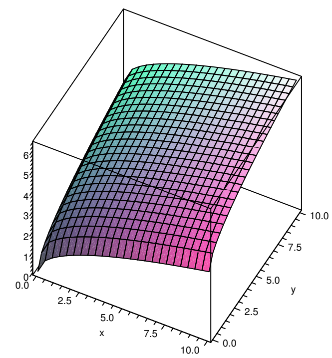


♣

Example 4.5: Square Root

Describe and graph the function $f(x,y)=\sqrt{x}+\sqrt{y}$.

Solution. This function is defined only when both x and y are non-negative. When $y=0$ we get $f(x,y)=\sqrt{x}$, the familiar square root function in the x-z-plane, and when $x=0$ we get the same curve in the y-z-plane. Generally speaking, we see that starting from $f(0,0)=0$ this function gets larger in every direction in roughly the same way that the square root function gets larger. For example, if we restrict attention to the line $x=y$, we get $f(x,y)=2\sqrt{x}$ and along the line $y=2x$ we have $f(x,y)=\sqrt{x}+\sqrt{2x}=(1+\sqrt{2})\sqrt{x}$ (see graph below).



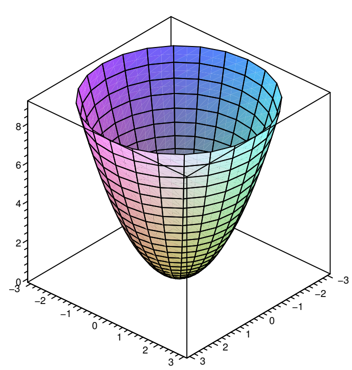
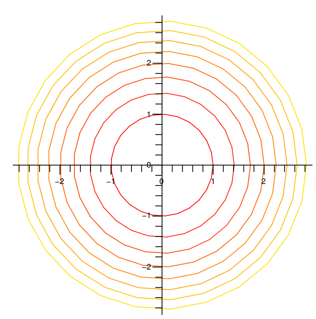
♣

Example 4.6: Elliptic Paraboloid

Describe and graph the function $f(x,y) = x^2 + y^2$.

Solution. When $x = 0$ this becomes $f(y) = y^2$, a parabola in the y-z-plane; when $y = 0$ we get the "same" parabola $f(x) = x^2$ in the x-z-plane.

Finally, picking a value $z = k$, at what points does $f(x,y) = k$? This means $x^2 + y^2 = k$, which we recognize as the equation of a circle of radius $\sqrt{k}$, as seen in the graph below to the right. So the graph of $f(x,y)$ has parabolic cross-sections as shown in the graph below to the left, and the same height everywhere on concentric circles with centre at the origin. This fits with what we have already discovered.

♣

As in this example, the points (x,y) such that $f(x,y) = k$ form a curve, called a **level curve** of the function. A graph of some level curves can give a good idea of the shape of the surface; it looks much like a topographic map of the surface. By drawing the level curves corresponding to several admissible values of k, we obtain a **contour map**. In the graph of Example 4.6 both the surface and its associated level curves are shown. Note that, as with a topographic map, the heights corresponding to the level curves are evenly spaced, so that where curves are closer together the surface is steeper.

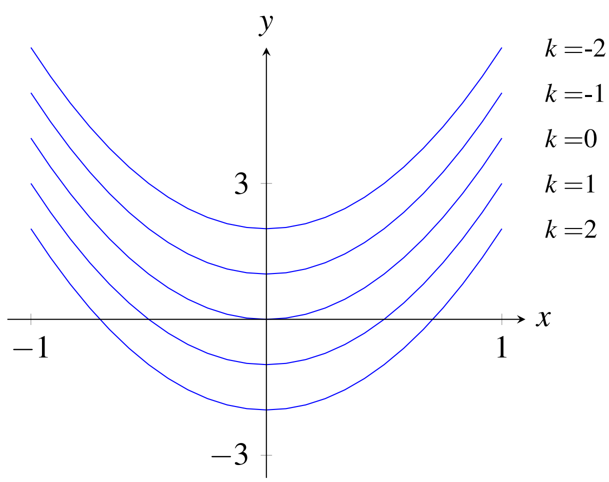
Example 4.7: Level Curves I

Sketch the level curves of the function $z = f(x,y) = 4x^2 - y$ corresponding to $z = -2, -1, 0, 1, 2$.

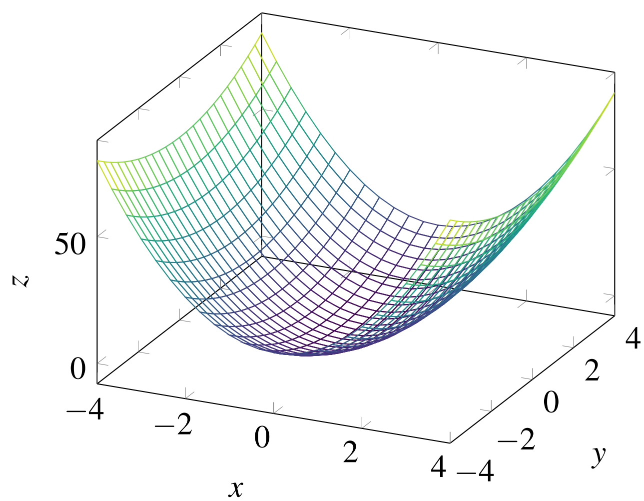
Solution. We find the level curves of f by setting $f(x,y) = z$ to be a constant. For $z = -2, -1, 0, 1, 2$, we find the equations

$$\begin{aligned} y &= 4x^2 + 2 \\ y &= 4x^2 + 1 \\ y &= 4x^2 \\ y &= 4x^2 - 1 \\ y &= 4x^2 - 2 \end{aligned}$$

and so each level curve is a parabola in the x-y-plane, where only the y-intercept changes. These level curves are shown in the graph below to the left, and the graph of $f(x,y)$ is shown to the right for reference.



(a) *Level curves of f*



(b) *The graph of $z = f(x,y)$*

♣

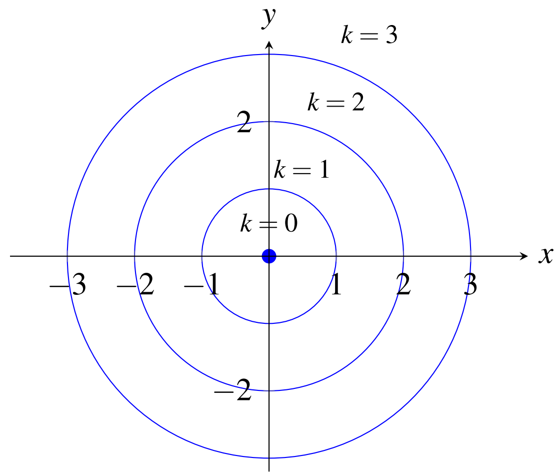
Example 4.8: Level Curves II

Sketch the level curves of the function $z = f(x,y) = \sqrt{x^2 + y^2}$ corresponding to $z = 0, 1, 2, 3$.

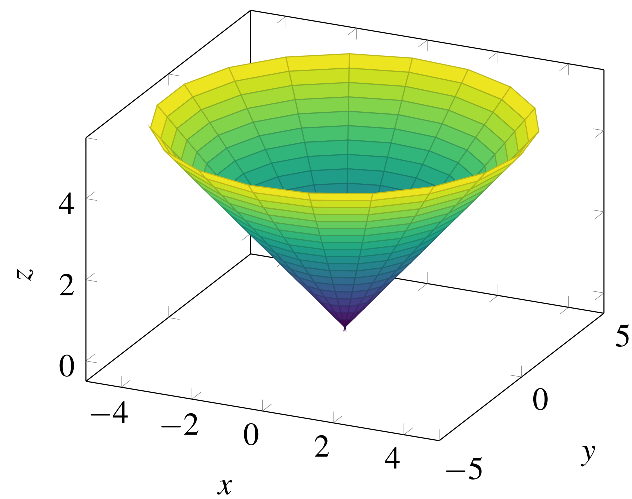
Solution. For each level curve corresponding to $z = c$, we have

$$\sqrt{x^2 + y^2} = c \implies x^2 + y^2 = c^2,$$

which is the equation of a circle centred at $(0,0)$ of radius c. These level curves are drawn below in the graph to the left. The graph of $z = f(x,y)$ is provided to the right.



(a) *Level curves of f*



(b) *The graph of $z = f(x,y)$*

♣

Example 4.9: Home Mortgage Payments

The monthly payment for a condo that amortizes a loan of A dollars in t years when the interest rate is r per year is given by

$$P = f(A,r,t) = \frac{Ar}{12\left[1-\left(1+\frac{r}{12}\right)^{-12t}\right]}$$

Find the monthly payment for a home mortgage of \$240,000 to be amortized over 25 years when the interest rate is 4% per year.

Solution. P is a function of three variables: A, r and t. To find the required monthly payments, we evaluate P at the given values,

$$\begin{aligned} P &= f(240000, 0.04, 25) \\ &= \frac{240000(0.04)}{12\left[1-\left(1+\frac{0.04}{12}\right)^{-12(25)}\right]} \\ &\approx 1266.81. \end{aligned}$$

Thus, the monthly payments would be about \$1266.81. ♣

Functions $f\colon \mathbb{R}^n \to \mathbb{R}$ for $n > 2$ behave much like functions of two variables. The principal difficulty with such functions is visualizing them, as they do not "fit" in the three dimensions we are familiar with. For $n = 3$ variables there are various ways to interpret functions that make them easier to understand. For example, $f(x,y,z)$ could represent the temperature at the point (x,y,z), or the strength of a magnetic field. Similar to considering level curves with two-variable functions, it is useful to consider those points at which $f(x_1,x_2,\ldots,x_n) = k$, where k is some constant for n-variable functions. This collection of points is called a **level set**. For three variables, a level set is a surface, called a **level surface**.

Example 4.10: Level Surfaces

Suppose the temperature at (x,y,z) *is* $T(x,y,z) = e^{-(x^2+y^2+z^2)}$. *Describe the level surfaces.*

Solution. This function has a maximum value of 1 at the origin, and tends to 0 in all directions. If k is positive and at most 1, the set of points for which $T(x,y,z) = k$ is those points satisfying $x^2 + y^2 + z^2 = -\ln k$, a sphere centred at the origin. The level surfaces are the concentric spheres centred at the origin. ♣

Exercises for Section 4.1

Exercise 4.1.1 *Find the equations and describe the shapes of the cross-sections when* $x = 0$, $y = 0$ *and* $x = y$. *Plot the surface with a three-dimensional graphing tool.*

(a) $f(x,y) = (x-y)^2$

(b) $f(x,y) = |x| + |y|$

(c) $f(x,y) = e^{-(x^2+y^2)} \sin(x^2 + y^2)$

(d) $f(x,y) = \sin(x-y)$

(e) $f(x,y) = (x^2 - y^2)^2$

Exercise 4.1.2 *Find the domain of each of the following functions:*

(a) $f(x,y) = \sqrt{9-x^2} + \sqrt{y^2-4}$

(b) $h(u,v) = \sqrt{2-u^2-v^2}$

(c) $f(x,y) = \sqrt{16-x^2-4y^2}$

(d) $f(x,y) = x + 4y$

(e) $g(x,y,z) = x^2 + y^2 + z^2$

(f) $f(x,y) = \arcsin(x^2 + y^2 - 2)$

(g) $h(u,v) = \dfrac{uv}{u-v}$

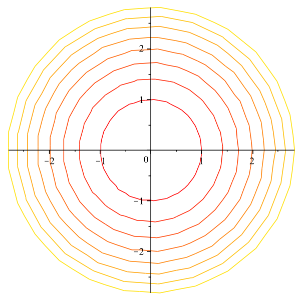
(h) $f(s,t) = \sqrt{s^2+t^2}$

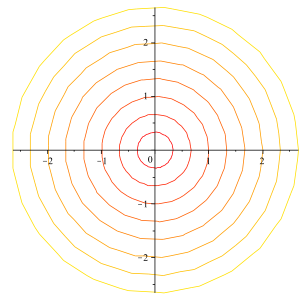
(i) $g(r,s) = \sqrt{rs}$

(j) $f(x,y) = e^{-xy}$

(k) $h(x,y) = \ln(x + y - 7)$

Exercise 4.1.3 *Below are two sets of level curves at a sequence of equally-spaced heights z. One is for a* *cone, one is for a paraboloid. Which is which? Explain.*





Exercise 4.1.4 *Sketch the level curves of the function corresponding to the given values of z.*

(a) $f(x,y) = 3x+2y$, $z = -2,-1,0,1,2$

(b) $f(x,y) = x^2 - y$, $z = -2,-1,0,1,2$

(c) $f(x,y) = 2y+x^2$, $z = -2,-1,0,1,2$

(d) $f(x,y) = 2xy$, $z = -4,-2,2,4$

(e) $f(x,y) = \sqrt{25-x^2-y^2}$, $z = 0,1,3,5$

(f) $f(x,y) = e^y - x^2$, $z = 1,2,3,4$

Exercise 4.1.5 *A clothing retailer sells casual and business jackets. Let q_c denote the quantity of casual jackets demanded monthly and q_b denote the quantity of business jackets demanded monthly. It is determined that the unit price of casual jackets is approximately*

$$p_c = 100 - \frac{1}{4}q_c - \frac{1}{7}q_b,$$

and that the unit price of business jackets is approximately

$$p_b = 150 - \frac{1}{10}q_c - \frac{1}{3}q_b$$

dollars, respectively.

(a) *Determine the monthly revenue, $R(q_c,q_b)$.*

(b) *Determine the domain of R.*

Exercise 4.1.6 *A certain manufacturer produces basic and enhanced versions of their product. Let q_b denote the quantity of basic units daily demanded and q_e denote the quantity of enhanced units demanded daily. It is determined that the unit price of the basic units is approximately*

$$p_b = 10 - 0.1q_b - 0.5q_e,$$

and that the unit price of the enhanced units is approximately

$$p_e = 30 - 0.4q_b - q_e$$

dollars, respectively.

(a) *Determine the daily revenue, $R(q_b,q_e)$.*

(b) Determine the domain of R.

Exercise 4.1.7 *We can calculate the outstanding principal on a loan at the end of i months by the formula*

$$B(A,r,t,i) = A\left(\frac{\left(1+\frac{r}{12}\right)^{i}-1}{\left(1+\frac{r}{12}\right)^{12t}-1}\right) \quad 0 \leq i \leq 12t$$

where A is the principal loan, r is the annual interest rate, and t is the amortization period in years. Suppose the original amount borrowed is \$100,000, and the interest rate charged by the bank is 3%. What is the amount owed to the bank after 2 years if the loan is to be repaid in equal instalments over 25 years? What is the amount owed after 20 years?

Exercise 4.1.8 *In economics, the given optimal quantity Q of goods for a store to order is given by the Wilson lot-size formula:*

$$Q(C,N,H) = \sqrt{\frac{2CN}{H}}$$

where C is the cost of placing an order, N is the number of items the store sells per day, and H is the daily holding cost for each item. What is the most economical quantity of winter tires to order if the store pays \$25 for placing an order, pays \$2 for holding a tire per day, and expects to sell 60 tires a day?

4.2 Double Integrals: Volume and Average Value

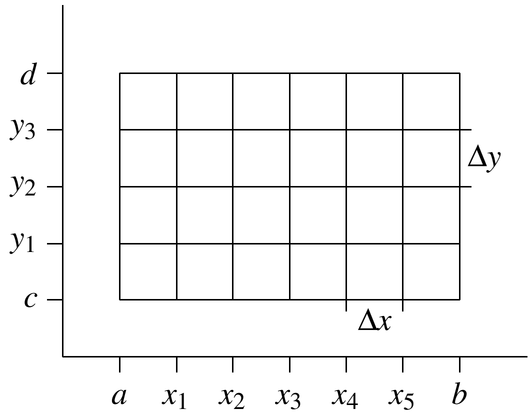
Consider a surface $f(x,y)$. In this section, we are interested in computing either the volume under f or the average function value of f over a certain area in the x-y-plane. You might temporarily think of this surface as representing physical topography—a hilly landscape, perhaps. What is the average height of the surface (or average altitude of the landscape) over some region?

4.2.1. Volume and Average Value over a Rectangular Region

As with most such problems, we start by thinking about how we might approximate the answer. Suppose the region is a rectangle, $[a,b] \times [c,d]$. We can divide the rectangle into a grid, m subdivisions in one direction and n in the other, as indicated in the 2D graph below, where $m = 6$ and $n = 4$. We pick x-values x_0, $x_1, \ldots, x_{m-1}$ in each subdivision in the x-direction, and similarly in the y-direction. At each of the points (x_i, y_j) in one of the smaller rectangles in the grid, we compute the height of the surface: $f(x_i, y_j)$. Now the average of these heights should be (depending on the fineness of the grid) close to the average height of the surface:

$$\frac{f(x_0,y_0)+f(x_1,y_0)+\cdots+f(x_0,y_1)+f(x_1,y_1)+\cdots+f(x_{m-1},y_{n-1})}{mn}$$

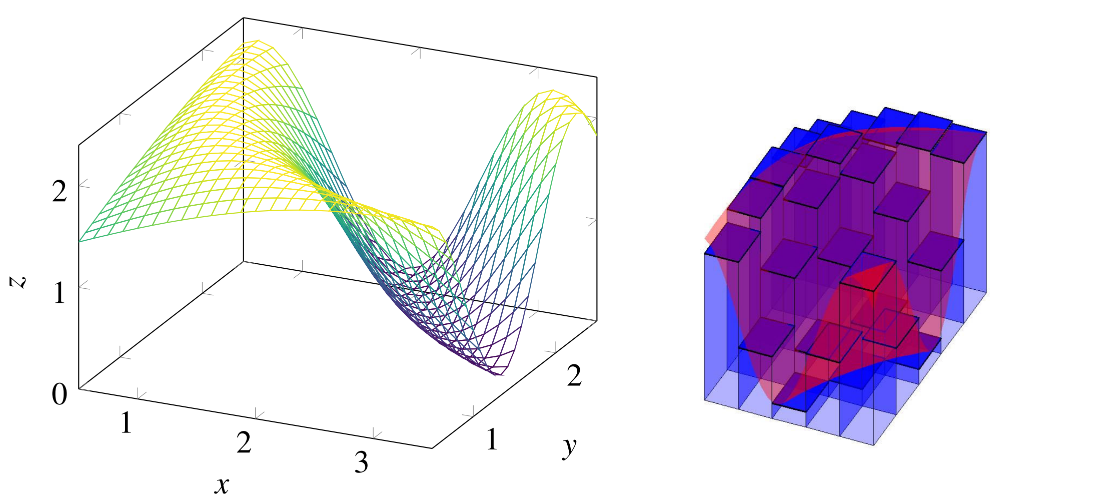
As both m and n go to infinity, we expect this approximation to converge to a fixed value, the actual average height of the surface. For reasonably nice functions this does indeed happen.



Using sigma notation, we can rewrite the approximation:

$$\begin{aligned}\frac{1}{mn}\sum_{i=0}^{n-1}\sum_{j=0}^{m-1} f(x_j,y_i) &= \frac{1}{(b-a)(d-c)}\sum_{i=0}^{n-1}\sum_{j=0}^{m-1} f(x_j,y_i)\frac{b-a}{m}\frac{d-c}{n} \\ &= \frac{1}{(b-a)(d-c)}\sum_{i=0}^{n-1}\sum_{j=0}^{m-1} f(x_j,y_i)\Delta x\Delta y.\end{aligned}$$

The two parts of this product have useful meaning: $(b-a)(d-c)$ is of course the area of the rectangle, and the double sum adds up mn terms of the form $f(x_j,y_i)\Delta x\Delta y$, which is the height of the surface at a point multiplied by the area of one of the small rectangles into which we have divided the large rectangle. In short, each term $f(x_j,y_i)\Delta x\Delta y$ is the volume of a tall, thin, rectangular box, and is approximately the volume under the surface and above one of the small rectangles; see Figure 4.1. When we add all of these up, we get an approximation to the volume under the surface and above the rectangle $R=[a,b]\times[c,d]$. When we take the limit as m and n go to infinity, the double sum becomes the actual volume under the surface, which we divide by $(b-a)(d-c)$ to get the average height.



(a) *The surface z on* $[0.5,3.5]\times[0.5,2.5]$. (b) *Approximating the volume under z.*

Figure 4.1: The surface $z=\sin(xy)+\frac{6}{5}$.

Double sums like this come up in many applications, so in a way it is the most important part of this example; dividing by $(b-a)(d-c)$ is a simple extra step that allows the computation of an average. This computation is independent of f being positive, and so as we did in the single variable case, we introduce a special notation for the limit of such a double sum, which is referred to as the **double integral** of f over the region R.

Definition 4.11: Double Integral over a Rectangular Region

Let $f(x,y)$ be a continuous function defined over the rectangle $R = [a,b] \times [c,d]$, then the **double integral** *of f over R is denoted by*

$$\iint_R f(x,y)\,dA = \iint_R f(x,y)\,dx\,dy = \lim_{m\to\infty}\lim_{n\to\infty}\sum_{i=0}^{n-1}\sum_{j=0}^{m-1} f(x_j,y_i)\Delta x\Delta y,$$

provided the limit exists.

Note: The notation dA indicates a small element of area, without specifying any particular order for the variables x and y; it is shorter and more generic than writing $dx\,dy$.

We now capture our results from the earlier calculations using the notation of the double integral.

Theorem 4.12: Average Value of a Two-variable Function

Let $f(x,y)$ be a continuous function defined over the rectangle $R = [a,b] \times [c,d]$, then the **average value** *f_{avg} of f over R is*

$$f_{avg} = \frac{1}{(b-a)(d-c)} \iint_R f(x,y)\,dA,$$

provided the double integral exists.

Theorem 4.13: Volume Below a Surface

Let $f(x,y)$ be a continuous function defined over the rectangle $R = [a,b] \times [c,d]$, with $f(x,y) \geq 0$, then the **volume** *V* **below the surface** *f is*

$$V = \iint_R f(x,y)\,dA,$$

provided the double integral exists.

Note: Just as with single-variable integration, this last theorem can be stated for more general functions that can take on negative values, as long as we understand that V represents a signed volume.

4.2.2. Computing Double Integrals over Rectangular Regions

The next question, of course, is: How do we compute these double integrals? You might think that we will need some two-dimensional version of the Fundamental Theorem of Calculus, but as it turns out we can get away with just the single variable version, applied twice.

Going back to the double sum, we can rewrite it to emphasize a particular order in which we want to add the terms:

$$\sum_{i=0}^{n-1}\left(\sum_{j=0}^{m-1} f(x_j,y_i)\Delta x\right)\Delta y.$$

In the sum in parentheses, only the value of x_j is changing; y_i is temporarily constant. As m goes to infinity, this sum has the right form to turn into an integral:

$$\lim_{m\to\infty}\sum_{j=0}^{m-1} f(x_j,y_i)\Delta x = \int_a^b f(x,y_i)\,dx.$$

So after we take the limit as m goes to infinity, the sum is

$$\sum_{i=0}^{n-1}\left(\int_a^b f(x,y_i)\,dx\right)\Delta y.$$

Of course, for different values of y_i this integral has different values; in other words, it is really a function applied to y_i:

$$G(y) = \int_a^b f(x,y)\,dx.$$

If we substitute back into the sum we get

$$\sum_{i=0}^{n-1} G(y_i)\Delta y.$$

This sum has a nice interpretation. The value $G(y_i)$ is the area of a cross section of the region under the surface $f(x,y)$, namely, when $y = y_i$. The quantity $G(y_i)\Delta y$ can be interpreted as the volume of a solid with face area $G(y_i)$ and thickness Δy. Think of the surface $f(x,y)$ as the top of a loaf of sliced bread. Each slice has a cross-sectional area and a thickness; $G(y_i)\Delta y$ corresponds to the volume of a single slice of bread. Adding these up approximates the total volume of the loaf. (This is very similar to the technique we used to compute volumes in Section 3.3, except that there we need the cross-sections to be in some way "the same".) Figure 4.2 shows this "sliced loaf" approximation using the same surface as shown in Figure 4.1. Nicely enough, this sum looks just like the sort of sum that turns into an integral, namely,

$$\begin{aligned}\lim_{n\to\infty}\sum_{i=0}^{n-1} G(y_i)\Delta y &= \int_c^d G(y)\,dy\\ &= \int_c^d\int_a^b f(x,y)\,dx\,dy.\end{aligned}$$

Let's be clear about what this means: we first will compute the inner integral, temporarily treating y as a constant. We will do this by finding an antiderivative with respect to x, then substituting $x = a$ and $x = b$ and subtracting, as usual. The result will be an expression with no x variable but some occurrences of y. Then the outer integral will be an ordinary one-variable problem, with y as the variable.

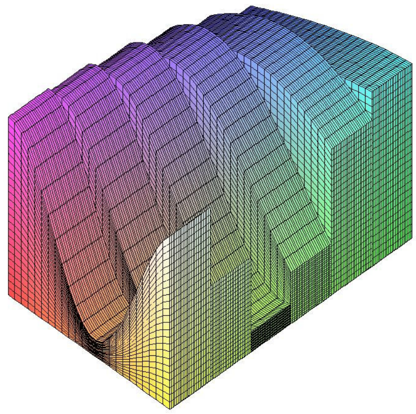


Figure 4.2: Approximating the volume under a surface with slices.

Example 4.14: Average Height

Figure 4.1 shows the function $\sin(xy) + 6/5$ *on* $[0.5, 3.5] \times [0.5, 2.5]$. *Find the average height of this surface.*

Solution. The volume under this surface is

$$\int_{0.5}^{2.5} \int_{0.5}^{3.5} \sin(xy) + \frac{6}{5}\,dx\,dy.$$

The inner integral is

$$\int_{0.5}^{3.5} \sin(xy) + \frac{6}{5}\,dx = \left.\frac{-\cos(xy)}{y} + \frac{6x}{5}\right|_{0.5}^{3.5} = \frac{-\cos(3.5y)}{y} + \frac{\cos(0.5y)}{y} + \frac{18}{5}.$$

Unfortunately, this gives a function for which we can't find a simple antiderivative. To complete the problem we could use Sage or similar software to approximate the integral. Doing this gives a volume of approximately 8.84, so the average height is approximately $8.84/6 \approx 1.47$. ♣

Because addition and multiplication are commutative and associative, we can rewrite the original double sum:

$$\sum_{i=0}^{n-1} \sum_{j=0}^{m-1} f(x_j, y_i)\Delta x \Delta y = \sum_{j=0}^{m-1} \sum_{i=0}^{n-1} f(x_j, y_i)\Delta y \Delta x.$$

Now if we repeat the development above, the inner sum turns into an integral:

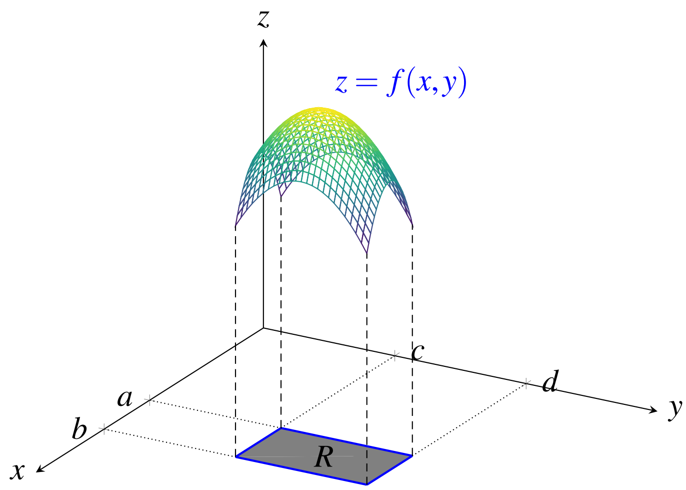
$$\lim_{n\to\infty} \sum_{i=0}^{n-1} f(x_j, y_i)\Delta y = \int_c^d f(x_j, y)\,dy,$$

and then the outer sum turns into an integral:

$$\lim_{m\to\infty} \sum_{j=0}^{m-1} \left(\int_c^d f(x_j, y)\,dy\right)\Delta x = \int_a^b \int_c^d f(x, y)\,dy\,dx.$$

In other words, we can compute the integrals in either order, first with respect to x then y, or vice versa. Thinking of the loaf of bread, this corresponds to slicing the loaf in a direction perpendicular to the first.

We summarize our findings by providing a general guideline for how the double integral over a rectangle, such as the one shown below, is computed.



Guideline for Computing a Double Integral over a Rectangle

Let $f(x,y)$ be a continuous function defined over the rectangle $R = [a,b] \times [c,d]$. Follow these steps to evaluate the double integral

$$\int_c^d \int_a^b f(x,y)\,dx\,dy = \int_a^b \int_c^d f(x,y)\,dy\,dx.$$

1. Choose the double integral, $\displaystyle\int_c^d \int_a^b f(x,y)\,dx\,dy$ or $\displaystyle\int_a^b \int_c^d f(x,y)\,dy\,dx$.
2. Compute the inner integral:
   (a) If $\displaystyle\int_a^b f(x,y)\,dx$, then temporarily treat y as a constant. Find an antiderivative with respect to x, then evaluate over the integration bounds $x = a$ and $x = b$.
   (b) If $\displaystyle\int_c^d f(x,y)\,dy$, then temporarily treat x as a constant. Find an antiderivative with respect to y, then evaluate over the integration bounds $y = c$ and $y = d$.
3. Compute the outer integral, which will be an ordinary one-variable integral of the form

$$\int_c^d G(y)\,dy \text{ or } \int_a^b G(x)\,dx.$$

We haven't really proved that the value of a double integral is equal to the value of the corresponding two single integrals in either order of integration, but provided the function is continuous, this is true. The result is called **Fubini's Theorem**, which we state here without proof.

Theorem 4.15: Fubini's Theorem for Rectangular Regions of Integration

Let $f(x,y)$ be a continuous function defined over the rectangle $R = [a,b] \times [c,d]$. Then the order of integration does not matter, and so

$$\iint_R f(x,y)\,dA = \int_c^d \int_a^b f(x,y)\,dx\,dy = \int_a^b \int_c^d f(x,y)\,dy\,dx.$$

We provide one example, where we compute the volume under a surface in two ways by switching the order of integration.

Example 4.16: Compute Volume in Two Ways

We compute $\iint_R 1+(x-1)^2+4y^2\,dA$, where $R = [0,3] \times [0,2]$, in two ways.

Solution. First,

$$\begin{aligned}
\int_0^3 \int_0^2 1+(x-1)^2+4y^2\,dy\,dx &= \int_0^3 y+(x-1)^2y+\frac{4}{3}y^3\Big|_0^2\,dx \\
&= \int_0^3 2+2(x-1)^2+\frac{32}{3}\,dx \\
&= 2x+\frac{2}{3}(x-1)^3+\frac{32}{3}x\Big|_0^3 \\
&= 6+\frac{2}{3}\cdot 8+\frac{32}{3}\cdot 3-(0-1\cdot\frac{2}{3}+0) \\
&= 44.
\end{aligned}$$

In the other order:

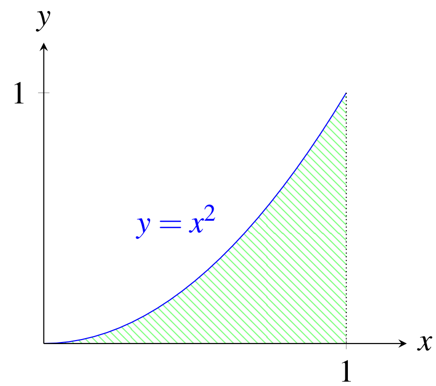
$$\begin{aligned}
\int_0^2 \int_0^3 1+(x-1)^2+4y^2\,dx\,dy &= \int_0^2 x+\frac{(x-1)^3}{3}+4y^2x\Big|_0^3\,dy \\
&= \int_0^2 3+\frac{8}{3}+12y^2+\frac{1}{3}\,dy \\
&= 3y+\frac{8}{3}y+4y^3+\frac{1}{3}y\Big|_0^2 \\
&= 6+\frac{16}{3}+32+\frac{2}{3} \\
&= 44.
\end{aligned}$$

♣

Note: In this example there is no particular reason to favor one direction over the other; in some cases, one direction might be much easier than the other, so it's usually worth considering the two different possibilities.

4.2.3. Computing Double Integrals over any Region

Frequently we will be interested in a region that is not simply a rectangle. Let's compute the volume under the surface $x+2y^2$ above the region described by $0 \le x \le 1$ and $0 \le y \le x^2$, shown below.



In principle there is nothing more difficult about this problem. If we imagine the three-dimensional region under the surface and above the parabolic region as an oddly shaped loaf of bread, we can still slice it up, approximate the volume of each slice, and add these volumes up. For example, if we slice perpendicular to the x-axis at x_i, the thickness of a slice will be Δx and the area of the slice will be

$$\int_0^{x_i^2} x_i + 2y^2\,dy.$$

When we add these up and take the limit as Δx goes to 0, we get the double integral

$$\begin{aligned}
\int_0^1 \int_0^{x^2} x+2y^2\,dy\,dx &= \int_0^1 xy + \frac{2}{3}y^3\Big|_0^{x^2}\,dx \\
&= \int_0^1 x^3 + \frac{2}{3}x^6\,dx \\
&= \frac{x^4}{4} + \frac{2}{21}x^7\Big|_0^1 \\
&= \frac{1}{4} + \frac{2}{21} = \frac{29}{84}.
\end{aligned}$$

We could just as well slice the solid perpendicular to the y-axis, in which case we get

$$\begin{aligned}
\int_0^1 \int_{\sqrt{y}}^1 x+2y^2\,dx\,dy &= \int_0^1 \frac{x^2}{2} + 2y^2x\Big|_{\sqrt{y}}^1\,dy \\
&= \int_0^1 \frac{1}{2} + 2y^2 - \frac{y}{2} - 2y^2\sqrt{y}\,dy \\
&= \frac{y}{2} + \frac{2}{3}y^3 - \frac{y^2}{4} - \frac{4}{7}y^{7/2}\Big|_0^1 \\
&= \frac{1}{2} + \frac{2}{3} - \frac{1}{4} - \frac{4}{7} = \frac{29}{84}.
\end{aligned}$$

What is the average height of the surface over this region? As before, it is the volume divided by the area of the base, but now we need to use integration to compute the area of the base, since it is not a simple rectangle. The area is

$$\int_0^1 x^2\,dx = \frac{1}{3},$$

so the average height is

$$\frac{29}{84} \div \frac{1}{3} = \frac{29}{28}.$$

Although we have not proven that the order of integration can be switched, we nonetheless capture our results and state the general version of Fubini's Theorem without proof.

Theorem 4.17: Fubini's Theorem for General Regions of Integration

Let $f(x,y)$ be a continuous function defined over the region R.

1. *If R is bounded by $a \le x \le b$, $g_1(x) \le y \le g_2(x)$, with g_1 and g_2 continuous on $[a,b]$, then*

$$\iint_R f(x,y)\,dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\,dy\,dx.$$

2. *If R is bounded by $c \le y \le d$, $h_1(y) \le x \le h_2(y)$, with h_1 and h_2 continuous on $[c,d]$, then*

$$\iint_R f(x,y)\,dA = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)\,dx\,dy.$$

Note: Although Fubini's Theorem tells us that the order of integration does not matter in a double integral, the theorem does not tell us which of the double integrals is easier to compute. Experience through practice allows us to decide whether to choose to set up a double integral with $dx\,dy$ or $dy\,dx$.

We explore the order of integration with one more example of a double integral.

Example 4.18: Volume of Region

Find the volume under the surface $z = \sqrt{1-x^2}$ and above the triangle formed by $y = x$, $x = 1$, and the x-axis.

Solution. Let's consider the two possible ways to set this up:

$$\int_0^1 \int_0^x \sqrt{1-x^2}\,dy\,dx \qquad \text{or} \qquad \int_0^1 \int_y^1 \sqrt{1-x^2}\,dx\,dy.$$

Which appears easier? In the first, the first (inner) integral is easy, because we need an antiderivative with respect to y, and the entire integrand $\sqrt{1-x^2}$ is constant with respect to y. Of course, the second integral may be more difficult. In the second, the first integral is mildly unpleasant—a trig substitution. So let's try the first one, since the first step is easy, and see where that leaves us.

$$\int_0^1 \int_0^x \sqrt{1-x^2}\,dy\,dx = \int_0^1 y\sqrt{1-x^2}\Big|_0^x\,dx = \int_0^1 x\sqrt{1-x^2}\,dx.$$

This is quite easy, since the substitution $u = 1 - x^2$ works:

$$\int x\sqrt{1-x^2}\,dx = -\frac{1}{2}\int \sqrt{u}\,du = \frac{1}{3}u^{3/2} = -\frac{1}{3}(1-x^2)^{3/2}.$$

Then

$$\int_0^1 x\sqrt{1-x^2}\,dx = -\frac{1}{3}(1-x^2)^{3/2}\Big|_0^1 = \frac{1}{3}.$$

This is a good example of how the order of integration can affect the complexity of the problem. In this case it is possible to do the other order, but it is a bit messier. In some cases one order may lead to an integral that is either difficult or impossible to evaluate; it's usually worth considering both possibilities before going very far. ♣

Exercises for Section 4.2

Exercise 4.2.1 *Compute the following double integrals.*

(a) $\int_0^2 \int_0^4 1 + x\,dy\,dx$

(b) $\int_{-1}^1 \int_0^2 x + y\,dy\,dx$

(c) $\int_1^2 \int_0^y xy\,dx\,dy$

(d) $\int_0^1 \int_{y^2/2}^{\sqrt{y}} dx\,dy$

(e) $\int_1^2 \int_1^x \frac{x^2}{y^2}\,dy\,dx$

(f) $\int_0^1 \int_0^{x^2} \frac{y}{e^x}\,dy\,dx$

(g) $\int_0^{\sqrt{\pi/2}} \int_0^{x^2} x\cos y\,dy\,dx$

(h) $\int_0^{\pi/2} \int_0^{\cos\theta} r^2(\cos\theta - r)\,dr\,d\theta$

(i) $\int_0^1 \int_{\sqrt{y}}^1 \sqrt{x^3+1}\,dx\,dy$

(j) $\int_0^1 \int_{y^2}^1$

(k) $\int_0^1 \int_{x^2}^1 x\sqrt{1+y^2}\,dy\,dx$

(l) $\int_0^1 \int_0^y \frac{2}{\sqrt{1-x^2}}\,dx\,dy$

(m) $\int_0^1 \int_{3y}^3 e^{x^2}\,dx\,dy$

(n) $\int_{-1}^1 \int_0^{1-x^2} x^2 - \sqrt{y}\,dy\,dx$

(o) $\int_0^{\sqrt{2}/2} \int_{-\sqrt{1-2x^2}}^{\sqrt{1-2x^2}} x\,dy\,dx$

Exercise 4.2.2 *Determine the volume of the region that is bounded as follows:*

(a) *below $z = 1 - y$ and above the region $-1 \le x \le 1$, $0 \le y \le 1 - x^2$*

(b) *bounded by $z = x^2 + y^2$ and $z = 4$*

(c) *in the first octant bounded by* $y^2 = 4 - x$ *and* $y = 2z$

(d) *in the first octant bounded by* $y^2 = 4x$, $2x + y = 4$, $z = y$, *and* $y = 0$

(e) *in the first octant bounded by* $x + y + z = 9$, $2x + 3y = 18$, *and* $x + 3y = 9$.

(f) *in the first octant bounded by* $x^2 + y^2 = a^2$ *and* $z = x + y$

(g) *bounded by* $4x^2 + y^2 = 4z$ *and* $z = 2$

(h) *bounded by* $z = x^2 + y^2$ *and* $z = y$

(i) *under the surface* $z = xy$ *and above the triangle with vertices* $(1,1,0)$, $(4,1,0)$, $(1,2,0)$

(j) *enclosed by* $y = x^2$, $y = 4$, $z = x^2$, $z = 0$

Exercise 4.2.3 *Evaluate* $\iint x^2\, dA$ *over the region in the first quadrant bounded by the hyperbola* $xy = 16$ *and the lines* $y = x$, $y = 0$, *and* $x = 8$.

Exercise 4.2.4 *A swimming pool is circular with a 40 meter diameter. The depth is constant along east-west lines and increases linearly from 2 meters at the south end to 7 meters at the north end. Find the volume of the pool.*

Exercise 4.2.5 *Find the average value of* $f(x,y) = e^y\sqrt{x + e^y}$ *on the rectangle with vertices* $(0,0)$, $(4,0)$, $(4,1)$ *and* $(0,1)$.

Exercise 4.2.6 *Reverse the order of integration on each of the following integrals*

(a) $\displaystyle\int_0^9 \int_0^{\sqrt{9-y}} f(x,y)\, dx\, dy$

(b) $\displaystyle\int_1^2 \int_0^{\ln x} f(x,y)\, dy\, dx$

(c) $\displaystyle\int_0^1 \int_{\arcsin y}^{\pi/2} f(x,y)\, dx\, dy$

(d) $\displaystyle\int_0^1 \int_{4x}^{4} f(x,y)\, dy\, dx$

(e) $\displaystyle\int_0^3 \int_0^{\sqrt{9-y^2}} f(x,y)\, dx\, dy$

4.3 Triple Integrals: Volume and Average Value

It will come as no surprise that we can also do triple integrals—integrals over a three-dimensional region. The simplest application allows us to compute volumes in an alternate way.

We follow the same method as we have done when we defined a single integral for functions of one variable and a double integral for functions of two variables. Suppose that $f(x,y,z)$ is a continuous function on a closed bounded region S in space. If the boundaries of S are "relatively smooth", then we can divide the three-dimensional region into small rectangular boxes with dimensions $\Delta x \times \Delta y \times \Delta z$ and with volume $dV = \Delta x \Delta y \Delta z$. Then we add them all up and take the limit, to get an integral:

$$\iiint_S f(x,y,z)\,dV.$$

Note:

1. Fubini's Theorem also holds for triple integrals, which means that the order of integration does not matter and we can choose from setting up a triple integral with any of the following six choices for the order of integration: $dx\,dy\,dz$, $dx\,dz\,dy$, $dy\,dx\,dz$, $dy\,dz\,dx$, $dz\,dx\,dy$, and $dz\,dy\,dx$. However, the same word of caution holds here as well, as some orders may lead to a more readily computable triple integral, while others may simply be too difficult to compute.

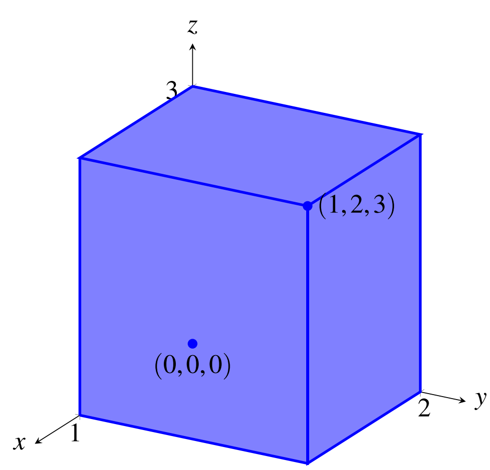
2. If the three-variable function f is the constant 1, then the triple integral $\iiint_S dV$ evaluates to the volume of the closed bounded region S.

3. If the three-variable function f is the constant 1 and S is bounded by constants, then we are simply computing the volume of a rectangular box.

Example 4.19: Volume of a Box

Compute the volume of the box with opposite corners at $(0,0,0)$ *and* $(1,2,3)$.

Solution. We begin by drawing an outline of the rectangular box as shown below.



Since the faces of the rectangular box are parallel to the coordinate planes, we deduce that the integration bounds are given by

$$0 \le x \le 1,\ 0 \le y \le 2,\ 0 \le z \le 3.$$

Hence, the following triple integral computes the volume of the rectangular box:

$$\int_0^1 \int_0^2 \int_0^3 dz\,dy\,dx = \int_0^1 \int_0^2 z\Big|_0^3 dy\,dx = \int_0^1 \int_0^2 3\,dy\,dx = \int_0^1 3y\Big|_0^2 dx = \int_0^1 6\,dx = 6.$$

Note that any of the following triple integrals would have resulted in the volume of the box:

$$\int_0^3\int_0^2\int_0^1 dx\,dy\,dz, \quad \int_0^3\int_0^2\int_0^1 dx\,dy\,dz, \quad \int_0^3\int_0^1\int_0^2 dy\,dx\,dz,$$

$$\int_0^1\int_0^3\int_0^2 dy\,dz\,dx, \quad \int_0^2\int_0^1\int_0^3 dz\,dx\,dy, \text{ or } \int_0^1\int_0^2\int_0^3 dz\,dy\,dx.$$

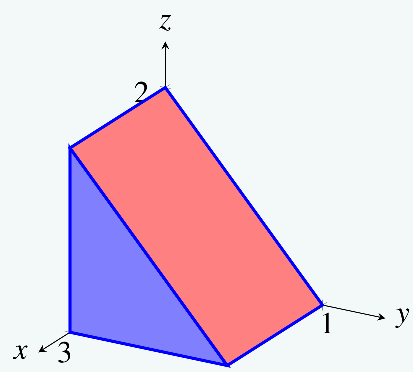
♣

Of course, this is more interesting and useful when the limits are not constant.

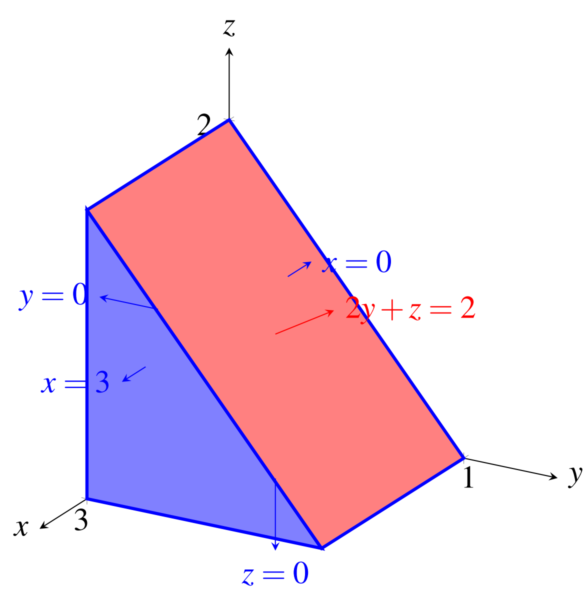
Example 4.20: Order of Integration

Calculate the volume of the prism shown using the order of integration

(a) $dx\,dy\,dz$ *(b)* $dy\,dz\,dx$

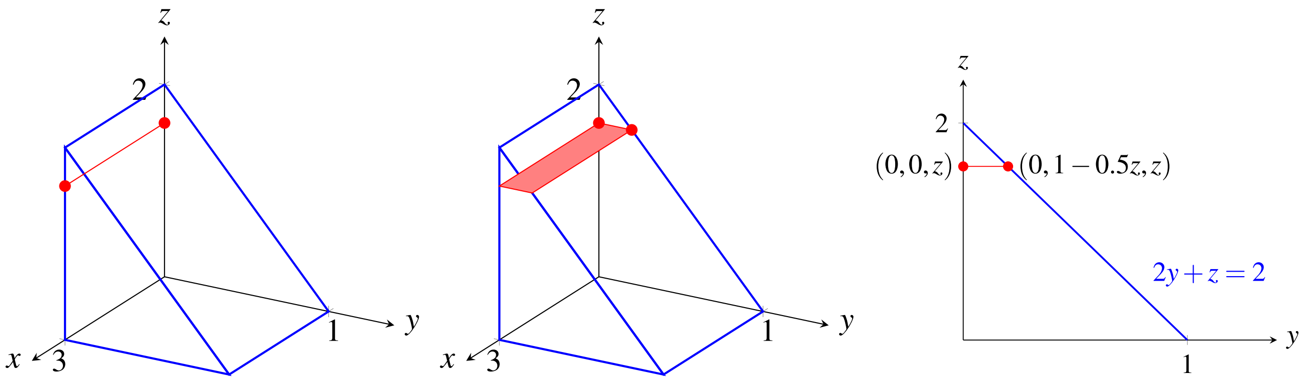


Solution. Let V be the volume of the prism. We begin by identifying the plane for each face of the prism:



(a) For the triple integral with order of integration $dx\,dy\,dz$, we begin by drawing a line parallel to the x-axis as shown below to the left that cuts through the prism. We notice that any such line stays between the values $x = 0$ and $x = 3$ and is thus constant. Hence, the inner most integral is

$$\int_?^? \int_?^? \int_0^3 dx\,dy\,dz.$$



Next, we are now extending the line in the x-y-plane to create a cross-sectional area that slices through the prism perpendicularly to the z-axis as shown above in the centre. These horizontal cross-sections are not constant and vary depending on the z-value that is chosen as shown above to the right. The z-values themselves are bounded by $z = 0$ and $z = 2$. Hence, the triple integral needed to evaluate the volume of the prism is

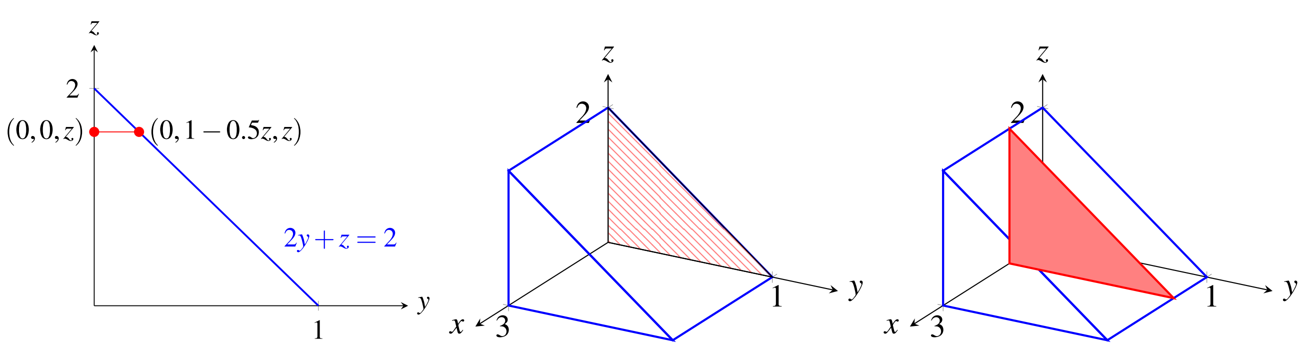
$$V = \int_0^2 \int_0^{1-0.5z} \int_0^3 dx\,dy\,dz.$$

Following through with the integration yields

$$\begin{aligned} V &= \int_0^2 \int_0^{1-0.5z} \int_0^3 dx\,dy\,dz = \int_0^2 \int_0^{1-0.5z} x\Big|_0^3 dy\,dz = \int_0^2 \int_0^{1-0.5z} 3\,dy\,dz \\ &= \int_0^2 3y\Big|_0^{1-0.5z} dz = \int_0^2 3(1-0.5z)\,dz = \left[3z - \frac{3}{4}z^2\right]_0^2 = 3. \end{aligned}$$

(b) For the triple integral with order of integration $dy\,dz\,dx$, we begin by drawing a line parallel to the y-axis as shown below to the left that cuts through the prism, since we are integrating with respect to y. We notice that any such line stays between the values $y = 0$ and $y = 1 - 0.5z$, and is thus variant depending on the z-value that is chosen. Hence, the inner most integral is

$$\int_?^? \int_?^? \int_0^{1-0.5z} dy\,dz\,dx.$$



Next, we integrate with respect to z, which means we are summing up all such lines from above that are perpendicular to the z-axis in the y-z-plane of the prism as shown above in the centre, which of course sums to the area of the triangle in the y-z-plane. These lines are bounded by $z = 0$ and $z = 2$. Hence, the inner two integrals needed to evaluate the volume of the prism are

$$\int_{?}^{?} \int_{0}^{2} \int_{0}^{1-0.5z} dy\,dz\,dx.$$

Lastly, we integrate with respect to x, which means we are summing up all such areas from above that are perpendicular to the x-axis. These areas are constant and are bounded by $x = 0$ and $x = 3$ as shown above to the right. Hence, the outer integral is

$$\int_{0}^{3} \int_{0}^{2} \int_{0}^{1-0.5z} dy\,dz\,dx.$$

Following through with the integration yields the volume V to be

$$V = \int_{0}^{3} \int_{0}^{2} \int_{0}^{1-0.5z} dy\,dz\,dx = \int_{0}^{3} \int_{0}^{2} y\Big|_{0}^{1-0.5z} dz\,dz = \int_{0}^{3} \int_{0}^{2} (1-0.5z)\,dz\,dx$$
$$= \int_{0}^{3} \left[z - \frac{z^2}{4}\right]_{0}^{2} dx = \int_{0}^{3} 1\,dx = x\Big|_{0}^{3} = 3,$$

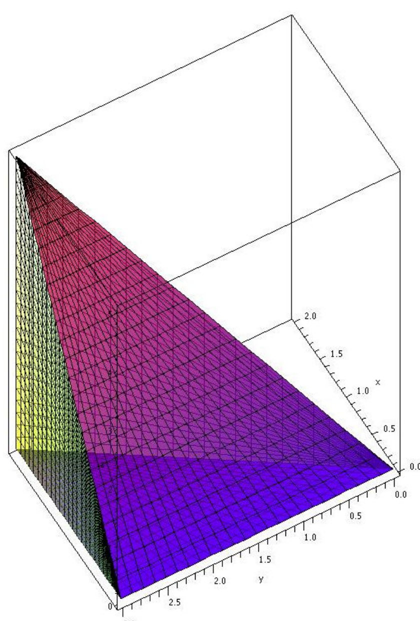
as we had before.

♣

Example 4.21: Volume of a Tetrahedron

Find the volume of the tetrahedron with corners at $(0,0,0)$, $(0,3,0)$, $(2,3,0)$, *and* $(2,3,5)$.

Solution. The whole problem comes down to correctly describing the region by inequalities: $0 \le x \le 2$, $3x/2 \le y \le 3$, $0 \le z \le 5x/2$. The lower y limit comes from the equation of the line $y = 3x/2$ that forms one edge of the tetrahedron in the x-y-plane; the upper z limit comes from the equation of the plane $z = 5x/2$ that forms the "upper" side of the tetrahedron as shown below.



Now the volume is

$$\begin{aligned}\int_0^2\int_{3x/2}^3\int_0^{5x/2} dz\,dy\,dx &= \int_0^2\int_{3x/2}^3 z\big|_0^{5x/2}\,dy\,dx\\ &= \int_0^2\int_{3x/2}^3 \frac{5x}{2}\,dy\,dx\\ &= \int_0^2 \frac{5x}{2}y\Big|_{3x/2}^3\,dx\\ &= \int_0^2 \frac{15x}{2}-\frac{15x^2}{4}\,dx\\ &= \frac{15x^2}{4}-\frac{15x^3}{12}\Big|_0^2\\ &= 15-10=5.\end{aligned}$$

♣

Pretty much just the way we did for two dimensions we can use triple integration in a variety of different physical, social and biological applications, and in computing various average quantities.

Example 4.22: Average Temperature in a Cube

Suppose the temperature at a point is given by $T = xyz$. Find the average temperature in the cube with opposite corners at $(0,0,0)$ and $(2,2,2)$.

Solution. In two dimensions:

1. Add up the temperature at each point in a region.
2. Divide by the area.

In three dimensions:

1. Add up the temperature at each point in space.
2. Divide by the volume.

Therefore, the average temperature in the cube is

$$\begin{aligned}\frac{1}{8}\int_0^2\int_0^2\int_0^2 xyz\,dz\,dy\,dx &= \frac{1}{8}\int_0^2\int_0^2 \frac{xyz^2}{2}\Big|_0^2\,dy\,dx\\ &= \frac{1}{16}\int_0^2\int_0^2 xy\,dy\,dx\\ &= \frac{1}{4}\int_0^2 \frac{xy^2}{2}\Big|_0^2\,dx\end{aligned}$$

$$= \frac{1}{8}\int_0^2 4x\,dx$$

$$= \frac{1}{2}\left.\frac{x^2}{2}\right|_0^2 = 1.$$

♣

Exercises for Section 4.3

Exercise 4.3.1 *Evaluate the following triple integrals.*

(a) $\displaystyle\int_0^1\int_0^x\int_0^{x+y}(2x+y-1)\,dz\,dy\,dx$

(b) $\displaystyle\int_0^2\int_{-1}^{x^2}\int_1^y xyz\,dz\,dy\,dx$

(c) $\displaystyle\int_0^1\int_0^x\int_0^{\ln y} e^{x+y+z}\,dz\,dy\,dx$

(d) $\displaystyle\int_0^{\pi/2}\int_0^{\sin\theta}\int_0^{r\cos\theta} r^2\,dz\,dr\,d\theta$

(e) $\displaystyle\int_0^{\pi}\int_0^{\sin\theta}\int_0^{r\sin\theta} r\cos^2\theta\,dz\,dr\,d\theta$

(f) $\displaystyle\int_0^1\int_0^{y^2}\int_0^{x+y} x\,dz\,dx\,dy$

(g) $\displaystyle\int_1^2\int_y^{y^2}\int_0^{\ln(y+z)} e^x\,dx\,dz\,dy$

(h) $\displaystyle\int_0^{\pi}\int_0^{\pi/2}\int_0^1 (z\sin x+z\cos y)\,dz\,dy\,dx$

Exercise 4.3.2 *Setup* $\iiint (x+y+z)\,dV$ *over the region inside* $x^2+y^2+z^2\le 1$ *in the first octant, but do not follow through on the integration as it requires a special technique that we have not introduced.*

Exercise 4.3.3 *Find the region E for which* $\iiint_E (1-x^2-y^2-z^2)\,dV$ *is a maximum.*

4.4 Probability

4.4.1. One Random Variable

4.4.1.1. Discrete Example

You perhaps have at least a rudimentary understanding of **discrete probability**, which measures the likelihood of an "event" when there are a finite number of possibilities. For example, when an ordinary six-sided die is rolled, the probability of getting any particular number is $1/6$. In general, the probability of an event is the number of ways the event can happen divided by the number of ways that "anything" can happen.

For a slightly more complicated example, consider the case of two six-sided dice. The dice are physically distinct, which means that rolling a 2–5 is different than rolling a 5–2; each is an equally likely event out of a total of 36 ways the dice can land, so each has a probability of $1/36$.

Most interesting events are not so simple. More interesting is the probability of rolling a certain sum out of the possibilities 2 through 12. It is clearly not true that all sums are equally likely: the only way to roll a 2 is to roll 1–1, while there are many ways to roll a 7. Because the number of possibilities is quite small, and because a pattern quickly becomes evident, it is easy to see that the probabilities of the various sums are:

$$\begin{aligned}
P(2) = \quad P(12) = &\quad 1/36\\
P(3) = \quad P(11) = &\quad 2/36\\
P(4) = \quad P(10) = &\quad 3/36\\
P(5) = \quad P(9) = &\quad 4/36\\
P(6) = \quad P(8) = &\quad 5/36\\
P(7) = &\quad 6/36
\end{aligned}$$

Here we use $P(n)$ to mean "the probability of rolling an n." Since we have correctly accounted for all possibilities, the sum of all these probabilities is $36/36 = 1$; the probability that the sum is one of 2 through 12 is 1, because there are no other possibilities.

The study of probability is concerned with more difficult questions as well; for example, suppose the two dice are rolled many times. On the average, what sum will come up? In the language of probability, this average is called the *expected value* of the sum. This is at first a little misleading, as it does not tell us what to "expect" when the two dice are rolled, but what we expect the long term average will be.

Suppose that two dice are rolled 36 million times. Based on the probabilities, we would expect about 1 million rolls to be 2, about 2 million to be 3, and so on, with a roll of 7 topping the list at about 6 million. The sum of all rolls would be 1 million times 2 plus 2 million times 3, and so on, and dividing by 36 million we would get the average:

$$\begin{aligned}
\bar{x} &= (2\cdot 10^6 + 3(2\cdot 10^6) + \cdots + 7(6\cdot 10^6) + \cdots + 12\cdot 10^6)\frac{1}{36\cdot 10^6}\\
&= 2\frac{10^6}{36\cdot 10^6} + 3\frac{2\cdot 10^6}{36\cdot 10^6} + \cdots + 7\frac{6\cdot 10^6}{36\cdot 10^6} + \cdots + 12\frac{10^6}{36\cdot 10^6}\\
&= 2P(2) + 3P(3) + \cdots + 7P(7) + \cdots + 12P(12)\\
&= \sum_{i=2}^{12} iP(i) = 7.
\end{aligned}$$

There is nothing special about the 36 million in this calculation. No matter what the number of rolls, once we simplify the average, we get the same $\sum_{i=2}^{12} iP(i)$. While the actual average value of a large number of rolls will not be exactly 7, the average should be close to 7 when the number of rolls is large. Turning this around, if the average is not close to 7, we should suspect that the dice are not fair.

4.4.1.2. Discrete and Continuous Random Variables

In this section we will introduce several concepts from probability concerning a single random variable for the purpose of showing yet another application of integration. In a subsequent section we extend the ideas presented here to showcase the use of double integrals.

Definition 4.23: Random Variable

A **random variable** X *is a variable that can take certain values, each with a corresponding probability.*

In the discrete example above, the random variable was the sum of the two dice.

Definition 4.24: Discrete Random Variable

When the number of possible values for X *is finite, we say that* X *is a* **discrete random variable** .

Definition 4.25: Continuous Random Variable

When the number of possible values for X *is infinite, we say that* X *is a* **continuous random variable**.

In many applications of probability, the number of possible values of a random variable is very large, perhaps even infinite. To deal with the infinite case we need a different approach, and since there is a sum involved, it should not be wholly surprising that integration turns out to be a useful tool. It then turns out that even when the number of possibilities is large but finite, it is frequently easier to pretend that the number is infinite. Suppose, for example, that a dart is thrown at a dart board. Since the dart board consists of a finite number of atoms, there are in some sense only a finite number of places for the dart to land, but it is easier to explore the probabilities involved by pretending that the dart can land on any point in the usual x-y-plane. Therefore, for the rest of this chapter we are concerned with continuous random variables.

4.4.1.3. Probability Density and Cumulative Distribution

Unlike for a discrete random variable, for a continuous random variable, we have that

$$P(X = x) = 0$$

for all x. We need to approach this differently and instead find the probability that X falls in some interval $[a,b]$. In other words, we need the density of probability of a continuous random variable, which defines the probability density function and allows us to calculate the probability that some value x of X falls in a given interval I.

Definition 4.26: Probability Density Function

Let f *be an integrable function. Then* f *is the* **probability density function** *of a continuous random variable* X *if* f *satisfies the following two properties:*

1. $f(x) \geq 0$ *for all* x.
2. $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1.$

Note:

1. We associate a probability density function with a random variable X by stipulating that the probability that X is between a and b is
$$P(a \leq X \leq b) = \int_a^b f(x)\,dx.$$
2. Since $P(X = x) = 0$ for all x, we have
$$P(a \leq x \leq b) = P(a < x \leq b) = P(a \leq x < b) = P(a < x < b).$$
3. Because of the requirement that the integral from $-\infty$ to ∞ be 1, all probabilities are less than or equal to 1, and the probability that X takes on some value between $-\infty$ and ∞ is 1, as it should be.

Example 4.27: Constructing a Probability Density Function

Construct a probability density function f from the following function g:
$$g(x) = \begin{cases} x^4 & x \in [0,1] \\ 0 & \text{all other } x. \end{cases}$$

Solution. First, a probability density function must be positive, and since $g(x) \geq 0$ for all x, this is true.

Second, we need that $\int_{-\infty}^{\infty} f(x)\,dx = 1$. Since
$$\int_{-\infty}^{\infty} x^4\,dx = \int_0^1 x^4\,dx = \left.\frac{x^5}{5}\right|_0^1 = \frac{1}{5},$$
we let
$$f(x) = \begin{cases} 5x^4 & x \in [0,1] \\ 0 & \text{all other } x. \end{cases}$$

♣

Example 4.28: Verifying a Probability Density Function I

Show that the following function f is a probability density function for $a < b$:
$$f(x) = \begin{cases} \dfrac{1}{b-a} & x \in [a,b] \\ 0 & \text{all other } x. \end{cases}$$

Solution. First, we need to show that f is positive:

Since $b < a$ we have that $b - a > 0$, and so $\frac{1}{b-a} > 0$. Thus, we have indeed that $f(x) \geq 0$.

Second, we verify that $\int_{-\infty}^{\infty} f(x)\,dx = 1$:

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)\,dx &= \int_a^b \frac{1}{b-a}\,dx \\ &= \frac{1}{b-a}x\Big|_a^b \\ &= \frac{b-a}{b-a} = 1.\end{aligned}$$

Hence, f is a probability density function. ♣

Note: The function in the above example is in fact the **uniform probability density function** (see Figure 4.3a).

Example 4.29: Verifying a Probability Density Function II

Show that the following function f is a probability density function for $c > 0$:

$$f(x) = \begin{cases} ce^{-cx} & x \in [0,\infty) \\ 0 & \text{all other } x. \end{cases}$$

Solution. First, we need to show that f is positive:

Since $c > 0$ and $e^{-cx} > 0$ for all x, we have indeed that $f(x) \geq 0$.

Second, we verify that $\int_{-\infty}^{\infty} f(x)\,dx = 1$:

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)\,dx &= \int_0^{\infty} ce^{-cx}\,dx - \lim_{a\to\infty}\int_0^a ce^{-cx}\,dx \\ &= \lim_{a\to\infty}\left[-e^{-cx}\right]_0^a \\ &= \lim_{a\to\infty}\left(-e^{-ca}\right) - (-1) = 1.\end{aligned}$$

Hence, f is a probability density function. ♣

Note: The function in the above example is in fact the **exponential probability density function** (see Figure 4.4a).

The entire collection of probabilities for a random variable X, namely $P(X \leq x)$ for all x, is called a **cumulative distribution**.

Definition 4.30: Cumulative Distribution Function

Suppose f is the probability density function of a random variable X. Then the cumulative distribution function is

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt.$$

Note:

1. For a continuous random variable X we have

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt,$$

and so taking the derivative with respect to x of both sides of the above equation, we see that

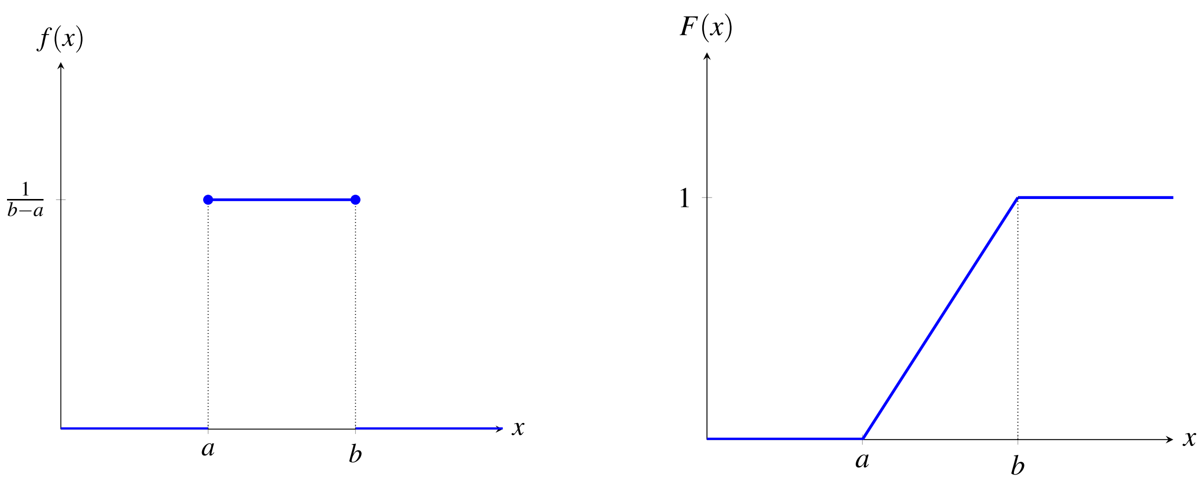
$$\frac{dF(x)}{dx} = f(x).$$

2. The **probability** that the random variable X belongs to an interval $[a,b] \subseteq \mathbb{R}$ is given by

$$P(a \leq X \leq b) = F(b) - F(a).$$

3. At times, the cumulative distribution function of a random variable X is written as $F_X(x)$.

We introduce three specific distribution functions, namely **uniform distribution**, or **rectangular distribution**, **exponential distribution** and **normal distribution**. In a uniform distribution (see Figure 4.3), all subintervals of equal length are equally probable.



(a) *Uniform probability density function.*

(b) *Uniform cumulative distribution function.*

Figure 4.3

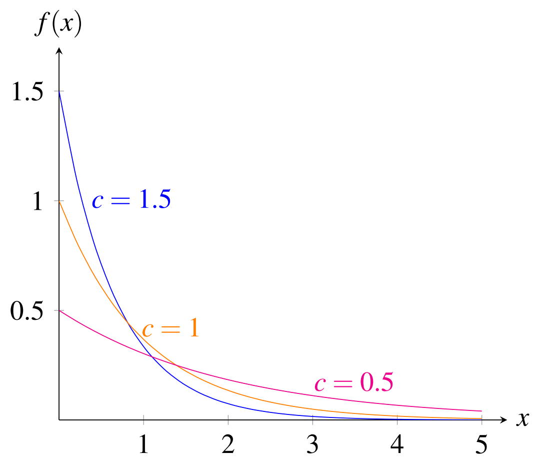
Definition 4.31: Uniform Distribution

Suppose that $a < b$ and

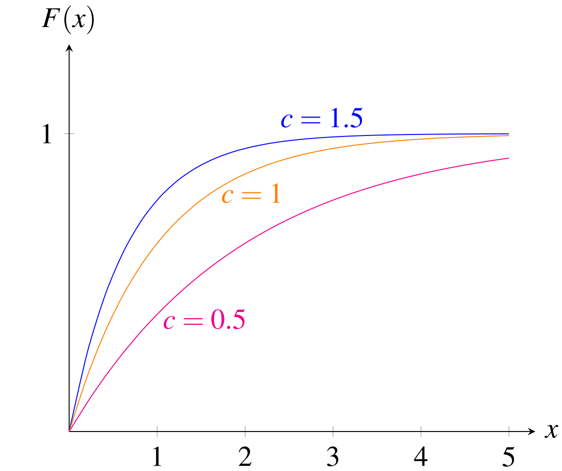
$$f(x) = \begin{cases} \dfrac{1}{b-a} & x \in [a,b] \\ 0 & \text{all other } x. \end{cases}$$

Then $f(x)$ is the **uniform probability density function** *on $[a,b]$ and the corresponding distribution is the* **uniform distribution** *on $[a,b]$.*

If the probability that an event occurs during a certain time interval is proportional to the length of that time interval–the wait time, then the wait time has exponential distribution (see Figure 4.4). This type of distribution allows us to answer questions such as "What is the wait time in a shopping queue?" or "What is the wait time making a call to some agency?".



(a) *Exponential probability density function.*



(b) *Exponential cumulative distribution function.*

Figure 4.4

Definition 4.32: Exponential Distribution

Suppose c is a positive constant and

$$f(x) = \begin{cases} 0 & x < 0 \\ ce^{-cx} & x \geq 0. \end{cases}$$

Then $f(x)$ is the **exponential probability density function** *and the corresponding distribution is the* **exponential distribution**.

Note: In the above definition, c is referred to as the **rate parameter** and is the constant of proportionality.

Example 4.33: Calculating Probabilities

Given the probability density function

$$f(x) = \begin{cases} 5x^4 & x \in [0,1] \\ 0 & \text{all other } x \end{cases}$$

of a continuous random variable X, calculate the following probabilities:

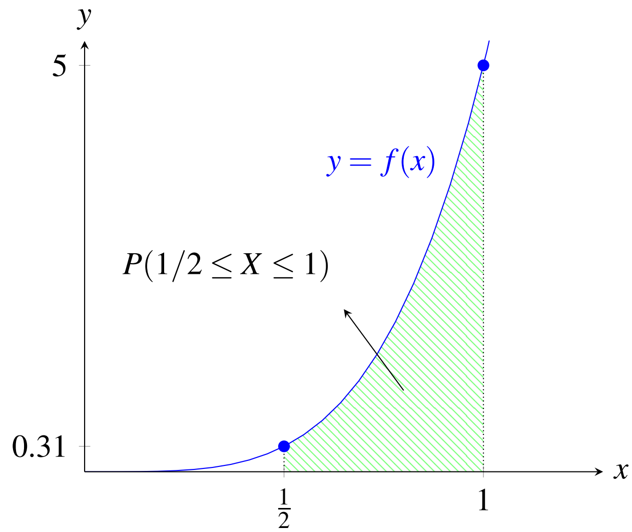
(a) $P(X = 1/2)$

(b) $P(1/2 \leq X \leq 1)$

Solution.

(a) $P(X=1/2)=\int_{1/2}^{1/2} 5x^4\,dx = x^5\big|_{1/2}^{1/2} = \left(\frac{1}{2}\right)^5 - \left(\frac{1}{2}\right)^5 = 0.$

(b) We begin by graphing f and the interested interval $[1/2, 1]$:



We are interested in the probability that X falls between $1/2$ and 1, which is the shaded area under the curve of f in the above graph:

$$P\left(\frac{1}{2} \le X \le 1\right) = \int_{1/2}^{1} 5x^4\,dx = x^5\Big|_{1/2}^{1} = 1^5 - \left(\frac{1}{2}\right)^5 = 0.96875.$$

♣

Example 4.34: Constructing a Special Probability Density Function

Consider the function $f(x) = e^{-x^2/2}$. What can we say about

$$\int_{-\infty}^{\infty} e^{-x^2/2}\,dx?$$

Use this information to construct a probability density function g from f.

Solution. First, it is easy to see that f is positive for all x. Next, we analyze

$$\int_{-\infty}^{\infty} e^{-x^2/2}\,dx.$$

We cannot find an antiderivative of f, but we can see that this integral is some finite number. Notice that $0 < f(x) = e^{-x^2/2} \le e^{-x/2}$ for $|x| > 1$. This implies that the area under $e^{-x^2/2}$ is less than the area under $e^{-x/2}$, over the interval $[1, \infty)$. It is easy to compute the latter area, namely

$$\int_{1}^{\infty} e^{-x/2}\,dx = \frac{2}{\sqrt{e}}.$$

By the Comparison Test, $\int_1^{\infty} e^{-x^2/2}\,dx$ is some finite number smaller than $2/\sqrt{e}$. Because f is symmetric around the y-axis,

$$\int_{-\infty}^{-1} e^{-x^2/2}\,dx = \int_1^{\infty} e^{-x^2/2}\,dx.$$

This means that

$$\int_{-\infty}^{\infty} e^{-x^2/2}\,dx = \int_{-\infty}^{-1} e^{-x^2/2}\,dx + \int_{-1}^{1} e^{-x^2/2}\,dx + \int_1^{\infty} e^{-x^2/2}\,dx = A$$

for some finite positive number A. Now if we let $g(x) = f(x)/A$,

$$\int_{-\infty}^{\infty} g(x)\,dx = \frac{1}{A}\int_{-\infty}^{\infty} e^{-x^2/2}\,dx = \frac{1}{A}A = 1,$$

so g is a probability density function.

We have shown that A is some finite number without computing it. By using some techniques from multi-variable calculus, it can be shown that $A = \sqrt{2\pi}$. ♣

The function

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

from the above example is the probability density function of a continuous random variable X which has a **standard normal distribution**. Before we say more about the standard normal distribution, let us introduce three more concepts, namely **expected value**, **variance** and **standard deviation**. Given a cumulative distribution, these concepts inform us about its central tendency and provide us with a measure of dispersion from this central tendency.

4.4.1.4. Expected Value, Variance and Standard Deviation

We ended our earlier discussion about the sum of two dice with a brief analysis of the **average** of such a sum. In probability, the average is often referred to as the **mean** or the **expected value**. This quantity is essentially calculated as the weighted average of all possible values of a random variable based on their probabilities. This means that if more and more values of a random variable were collected by repeated trials of a probability activity, then the sample mean becomes closer to the expected value, and as such, the expected value is the *long-run mean* of a random variable. For example, you want to know how well you perform on a multiple choice exam if you guess all the answers. Then the expected value tells you how many questions you might get right. We now formally introduce this concept for a discrete random variable.

Definition 4.35: Expected Value for a Discrete Random Variable

Suppose X is a discrete random variable. Then the **expected value** *of X is*

$$E(X) = x_1P(x_1) + x_2P(x_2) + \cdots + x_nP(x_n) = \sum_{i=1}^{n} x_iP(x_i)$$

where x_i are the values of X and $P(x_i)$ are the associated probabilities.

It comes as no surprise that for the calculation of the expected value of a continuous random variable the sum is extended to an integral.

Definition 4.36: Expected Value for a Continuous Random Variable

Suppose X is a continuous random variable. Then the **expected value** *of X is*

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$$

provided the integral converges.

Note:

1. The expected value is often denoted by the Greek symbol μ (read "mu").
2. The expected value does not always exist.
3. The expected value is essentially a type of centrality measure as it indicates the *typical* value for a probability distribution.

Example 4.37: Expected Value of the Standard Normal Distribution

Calculate the expected value of the standard normal distribution, where the probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Solution. The expected value of the standard normal distribution is

$$E(X) = \int_{-\infty}^{\infty} x \frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx.$$

We compute the two halves:

$$\int_{-\infty}^{0} x \frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx = \lim_{a\to-\infty} -\frac{e^{-x^2/2}}{\sqrt{2\pi}}\bigg|_a^0 = -\frac{1}{\sqrt{2\pi}}$$

and

$$\int_{0}^{\infty} x \frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx = \lim_{a\to\infty} -\frac{e^{-x^2/2}}{\sqrt{2\pi}}\bigg|_0^a = \frac{1}{\sqrt{2\pi}}.$$

Hence,

$$E(X) = \int_{-\infty}^{\infty} x \frac{e^{-x^2/2}}{\sqrt{2\pi}} = -\frac{1}{\sqrt{2\pi}} + \frac{1}{\sqrt{2\pi}} = 0.$$

Therefore, the expected value of the standard normal distribution is zero. ♣

Example 4.38: Expected Value of the Uniform Distribution

Calculate the expected value of the uniform distribution, where the probability density function is

$$f(x) = \begin{cases} \dfrac{1}{b-a} & x \in [a,b] \\ 0 & \textit{all other } x \end{cases}$$

with $a < b$.

Solution. The expected value of the uniform distribution is

$$E(X) = \int_{-\infty}^{a} x \cdot 0\,dx + \int_{a}^{b} \frac{x}{b-a}\,dx + \int_{b}^{\infty} x \cdot 0\,dx = \int_{a}^{b} \frac{x}{a-b}\,dx.$$

Therefore,

$$E(X) = \int_{a}^{b} \frac{x}{a-b}\,dx = \frac{1}{a-b} \cdot \frac{x^2}{2}\bigg|_{a}^{b} = \frac{1}{2}\frac{b^2-a^2}{b-a} = \frac{a+b}{2}.$$

And so the expected value of the uniform distribution is half the length of the interval $[a,b]$. ♣

Example 4.39: Expected Value of the Exponential Distribution

Calculate the expected value of the exponential distribution, where the probability density function is

$$f(x) = \begin{cases} ce^{-cx} & x \in [0,\infty) \\ 0 & \textit{all other } x \end{cases}$$

for $c > 0$.

Solution. The expected value of the exponential distribution is

$$E(X) = \int_{-\infty}^{0} x \cdot 0\,dx + \int_{0}^{\infty} x \cdot ce^{-cx}\,dx = \int_{0}^{\infty} x \cdot ce^{-cx}\,dx.$$

We calculate the indefinite integral using Integration by Parts:

$$\int x \cdot ce^{-cx}\,dx = -xe^{-cx} + \int e^{-cx}\,dx = -xe^{-cx} - \frac{1}{c}e^{-cx}.$$

Thus,

$$E(X) = \int_{0}^{\infty} x \cdot ce^{-cx}\,dx = \lim_{a\to\infty}\left[-xe^{-cx} - \frac{1}{c}e^{-cx}\right]_{0}^{a} = \frac{1}{c}.$$

♣

While the expected value is very useful, it typically is not enough information to properly evaluate a situation. For example, suppose we could manufacture an 11-sided die, with the faces numbered 2 through 12 so that each face is equally likely to be down when the die is rolled. The value of a roll is the value on this lower face. Rolling the die gives the same range of values as rolling two ordinary dice, but now each value occurs with probability $1/11$. The expected value of a roll is

$$\frac{2}{11}+\frac{3}{11}+\cdots+\frac{12}{11}=7,$$

which is the same value as the expected value of the earlier two-dice experiment. Therefore, the expected value does not distinguish the two cases, though of course they are quite different.

If f is a probability density function for a random variable X, with expected value μ, we would like to measure how far a *typical* value of X is from μ. One way to measure this distance is $(X-\mu)^2$; we square the difference so as to measure all distances as positive. To get the typical such squared distance, we compute the mean, which is referred to as the **variance** of a discrete random variable. For two dice, for example, we get

$$(2-7)^2\frac{1}{36}+(3-7)^2\frac{2}{36}+\cdots+(7-7)^2\frac{6}{36}+\cdots(11-7)^2\frac{2}{36}+(12-7)^2\frac{1}{36}=\frac{35}{36}.$$

Because we squared the differences this does not directly measure the typical distance we seek; if we take the square root of this we do get such a measure, $\sqrt{35/36}\approx 2.42$. The square root of the variance is called the **standard deviation** and denoted by the Greek letter σ (read "sigma"). Doing the computation for the strange 11-sided die we get

$$(2-7)^2\frac{1}{11}+(3-7)^2\frac{1}{11}+\cdots+(7-7)^2\frac{1}{11}+\cdots(11-7)^2\frac{1}{11}+(12-7)^2\frac{1}{11}=10,$$

with square root approximately 3.16. Comparing 2.42 to 3.16 tells us that the two-dice rolls clump somewhat more closely near 7 than the rolls of the weird die, which of course we already knew because these examples are quite simple.

Definition 4.40: Variance of a Discrete Random Variable

Suppose X is a discrete random variable. Then the **variance** *of X is*

$$\begin{aligned}V(X)&=(x_i-\mu)^2P(x_1)+(x_2-\mu)^2P(x_2)+\cdots+(x_n-\mu)^2P(x_n)\\&=\sum_{i=1}^{n}(x_i-\mu)^2P(x_i)\\&=\sum_{i=1}^{n}x_i^2P(x_i)-\mu^2\end{aligned}$$

where x_i are values of X, $P(x_i)$ are the associated probabilities, and μ is the expected value of X.

Note: The last equality in the above definition is an application of the definition of the expected value, and we leave it to the reader to show this is true.

Definition 4.41: Standard Deviation of a Discrete Random Variable

Suppose X is a discrete random variable. Then the **standard deviation** *of X is*

$$\sigma = \sqrt{\sum_{i=1}^{n} (x_i - \mu)^2 P(x_i)} = \sqrt{V(x)}$$

where x_i are values of X, $P(x_i)$ are the associated probabilities, and V is the variance of X.

To perform the same computation for a probability density function the sum is replaced by an integral, just as in the computation of the mean.

Definition 4.42: Variance of a Continuous Random Variable

Suppose X is a continuous random variable with probability density function f and expected value μ. Then the **variance** *of X is*

$$V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx = \int_{-\infty}^{\infty} x^2 f(x)\, dx - \mu^2.$$

Note: Just as for Definition 4.40, the last equality in the above definition can be readily derived from the definition of the expected value, and we leave it to the reader to show this is true.

Definition 4.43: Standard Deviation of a Continuous Random Variable

Suppose X is a continuous random variable with probability density function f and variance V. Then the **standard deviation** *of X is*

$$\sigma(X) = \sqrt{V(X)}.$$

Note:

1. The variance V of X is the dispersion from the mean.
2. The calculation of the variance is based on the mean, and so
$$V(X) = E((X-\mu)^2) = E((X - E(X))^2).$$
3. The variance is the mean of a squared number, and so $V(X) \geq 0$.
4. The larger the distance $(X-\mu)^2$ is on average, the higher the variance.
5. The variance of a constant random variable is zero, since then $E(X) = X$.

Example 4.44: Standard Deviation of the Standard Normal Distribution

Calculate the standard deviation of the standard normal distribution, where the probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Solution. We begin by finding the variance:

$$V(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2}\, dx.$$

To compute the antiderivative, use Integration by Parts, with $u = x$ and $dv = xe^{-x^2/2}\, dx$. This gives

$$\int x^2 e^{-x^2/2}\, dx = -xe^{-x^2/2} + \int e^{-x^2/2}\, dx.$$

We cannot compute the new integral, but we know its value when the limits are $-\infty$ to ∞, from our discussion of the standard normal distribution in Example 4.34.

$$\begin{aligned}
V(X) &= \int_{-\infty}^{\infty} x^2 e^{-x^2/2}\, dx \\
&= \left[-\frac{1}{\sqrt{2\pi}} xe^{-x^2/2}\right]_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2}\, dx \\
&= \lim_{a\to-\infty} \left[-\frac{1}{\sqrt{2\pi}} xe^{-x^2/2}\right]_a^0 + \lim_{a\to\infty} \left[-\frac{1}{\sqrt{2\pi}} xe^{-x^2/2}\right]_0^a + \frac{1}{\sqrt{2\pi}} \int_{\infty}^{\infty} e^{-x^2/2}\, dx \\
&= 0 + 0 + \frac{1}{\sqrt{2\pi}} \sqrt{2\pi} = 1
\end{aligned}$$

Therefore, the standard deviation of the standard normal distribution is $\sigma = \sqrt{1} = 1$. ♣

Example 4.45: Standard Deviation of the Uniform Distribution

Calculate the standard deviation of the uniform distribution, where the probability density function is

$$f(x) = \begin{cases} \dfrac{1}{b-a} & x \in [a,b] \\ 0 & \text{all other } x \end{cases}$$

where $a < b$.

Solution. The mean of the uniform distribution is found in Example 4.38 to be

$$\mu = \frac{a+b}{2}.$$

We now calculate the variance:

$$\begin{aligned}
V(X) &= \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a}\,dx = \int_a^b \frac{x^2}{b-a}\,dx - \left(\frac{a+b}{2}\right)^2 \\
&= \left.\frac{x^3}{3(b-a)}\right|_a^b - \frac{(a+b)^2}{4} \\
&= \frac{b^3-a^3}{3(b-a)} - \frac{(a+b)^2}{4} \\
&= \frac{1}{12}(b-a)^2
\end{aligned}$$

Hence, the standard deviation of the uniform distribution over the interval $[a,b]$ is

$$\sigma(X) = \sqrt{\frac{(b-a)^2}{12}} = \frac{b-a}{\sqrt{12}}.$$

♣

Example 4.46: Standard Deviation of the Exponential Distribution

Calculate the standard deviation of the exponential distribution, where the probability density function is

$$f(x) = \begin{cases} ce^{-cx} & x \in [0,\infty) \\ 0 & \textit{all other } x \end{cases}$$

for $c > 0$.

Solution. Recall from Example 4.39 that $\mu = \frac{1}{c}$. So the variance is given by

$$V(X) = \int_0^\infty \left(x - \frac{1}{c}\right)^2 ce^{-cx}\,dx,$$

which we can calculate using Integration by Parts:

$$\begin{aligned}
c\int \left(x-\frac{1}{c}\right)^2 e^{-cx}\,dx &= -\left(x-\frac{1}{c}\right)^2 e^{-cx} + \int 2\left(x-\frac{1}{c}\right) e^{-cx}\,dx \\
&= -\left(x-\frac{1}{c}\right)^2 e^{-cx} - \frac{2\left(x-\frac{1}{c}\right)}{c} e^{-cx} + \frac{2}{c}\int e^{-cx}\,dx \\
&= -\left(x-\frac{1}{c}\right)^2 e^{-cx} - \frac{2\left(x-\frac{1}{c}\right)}{c} e^{-cx} - \frac{2}{c^2} e^{-cx}
\end{aligned}$$

Therefore,

$$\begin{aligned}
V(X) &= \int_0^\infty \left(x-\frac{1}{c}\right)^2 ce^{-cx}\,dx \\
&= \lim_{a\to\infty}\left[-\left(x-\frac{1}{c}\right)^2 e^{-cx} - \frac{2\left(x-\frac{1}{c}\right)}{c} e^{-cx} - \frac{2}{c^2} e^{-cx}\right]_0^a = \frac{1}{c^2}.
\end{aligned}$$

Hence, the standard deviation of the exponential distribution is $\sigma(X) = \sqrt{\frac{1}{c^2}} = \frac{1}{c}$. ♣

4.4.1.5. Normal Distribution

One of the most prominent distributions in probability is the so-called **normal distribution** or **bell-shaped distribution**; the special case where $\sigma = 1$ and $\mu = 0$ was discussed in Example 4.37. Many important data sets, such as exam grades or annual precipitation on the West Coast of BC, can be modelled by a normal distribution. The following is a list of the characteristics of such a distribution:

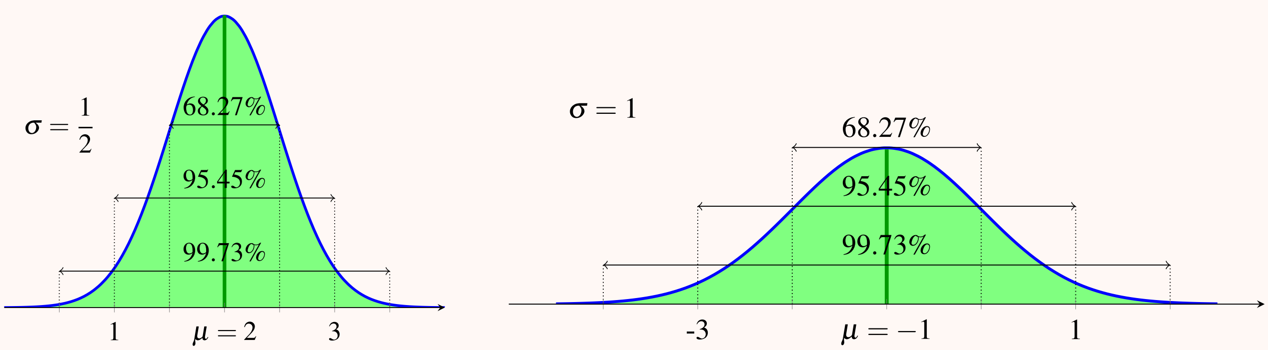
Characteristics of a Normal Distribution Function

Let X be a **normal random variable**, then its probability density function is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2},$$

where μ is the mean and σ is the standard deviation. The associated **normal distribution** has the following characteristics.

1. Its graph is a bell-shaped curve, and hence called **bell curve** (see sample graphs below).
2. The total area under the curve is 1.
3. The data are symmetrically distributed in the graph around its mean.
4. The data are concentrated around the mean.
5. The further a value is from the mean, the less probable it is to observe that value.
6. About 68.27% of the values are within one standard deviation of the mean.
   About 95.45% of the values are within two standard deviations of the mean.
   About 99.73% of the values – almost all of them – are within three standard deviations of the mean.



The **standard normal distribution** is the simplest case of a normal distribution, namely when $\mu = 0$ and $\sigma = 1$. The standard normal distribution allows us to compare distributions of data.

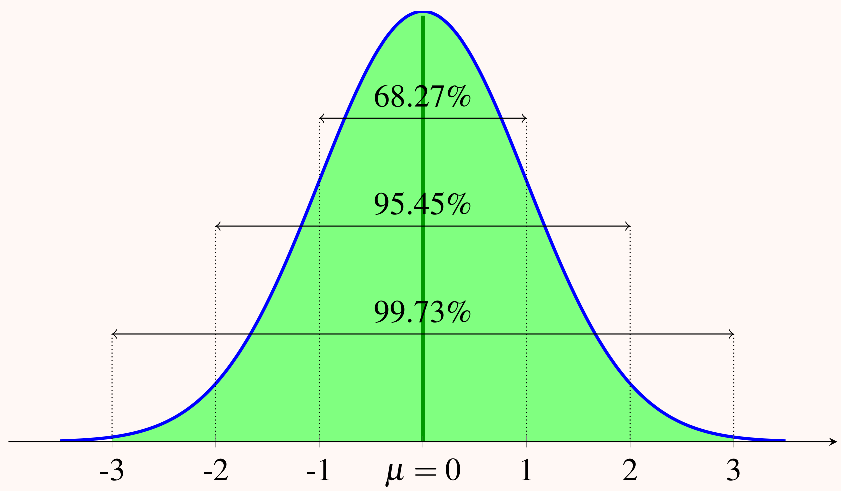
Characteristics of the Standard Normal Distribution Function

Let X be a **standard normal random variable**, then its probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

where μ is the mean and σ is the standard deviation. The associated **standard normal distribution** has the following characteristics.

1. The same characteristics as a Normal Distribution Function.
2. The mean is zero: $\mu = 0$.
3. The standard deviation is one: $\sigma = 1$.



Because normal distributions play such a vital role in statistics, the areas under the standard normal curve for any value of the normal random variable have been extensively computed and tabulated for easy reference. We will not concern ourselves with such tables. Instead, here is a simple example showing how these ideas can be useful.

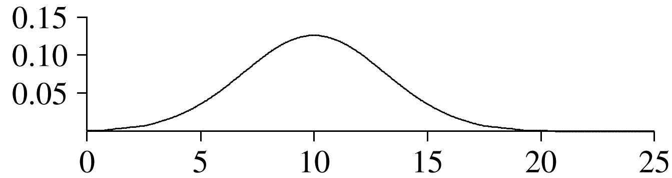
Example 4.47: Memory Chips

Suppose it is known that, in the long run, 1 out of every 100 computer memory chips produced by a certain manufacturing plant is defective when the manufacturing process is running correctly. Suppose 1000 chips are selected at random and 15 of them are defective. This is more than the expected number 10, but is it so many that we should suspect that something has gone wrong in the manufacturing process?

Solution. We are interested in the probability that various numbers of defective chips arise; the probability distribution is discrete: there can only be a whole number of defective chips. But (under reasonable assumptions) the distribution is very close to a normal distribution, namely this one:

$$f(x) = \frac{1}{\sqrt{2\pi}\sqrt{1000(.01)(.99)}} \exp\left(\frac{-(x-10)^2}{2(1000)(.01)(.99)}\right),$$

which is pictured below (recall that $\exp(x) = e^x$).



Now how do we measure how unlikely it is that under normal circumstances we would see 15 defective chips? We can't compute the probability of exactly 15 defective chips, as this would be $\int_{15}^{15} f(x)\,dx = 0$. We could compute $\int_{14.5}^{15.5} f(x)\,dx \approx 0.036$; this means there is only a 3.6% chance that the number of defective chips is 15. (We cannot compute these integrals exactly; computer software has been used to approximate the integral values in this discussion.) But this is misleading: $\int_{9.5}^{10.5} f(x)\,dx \approx 0.126$, which is larger, certainly, but still small, even for the "most likely" outcome. The most useful question, in most circumstances, is this: how likely is it that the number of defective chips is "far from" the mean? For example, how likely, or unlikely, is it that the number of defective chips is different by 5 or more from the expected value of 10? This is the probability that the number of defective chips is less than 5 or larger than 15, namely

$$\int_{-\infty}^{5} f(x)\,dx + \int_{15}^{\infty} f(x)\,dx \approx 0.11.$$

So there is an 11% chance that this happens—not large, but not tiny. Hence the 15 defective chips does not appear to be cause for alarm: about one time in nine we would expect to see the number of defective chips 5 or more away from the expected 10.

What if the observed number of defective chips was 20? Here we compute

$$\int_{-\infty}^{0} f(x)\,dx + \int_{20}^{\infty} f(x)\,dx \approx 0.0015.$$

So there is only a 0.15% chance that the number of defective chips is more than 10 away from the mean; this would typically be interpreted as too suspicious to ignore—it shouldn't happen if the process is running normally.

The big question, of course, is what level of improbability should trigger concern? It depends to some degree on the application, and in particular on the consequences of getting it wrong in one direction or the other. If we're wrong, do we lose a little money? A lot of money? Do people die? In general, the standard choices are 5% and 1%. So what we should do is find the number of defective chips that has only, let us say, a 1% chance of occurring under normal circumstances, and use that as the relevant number. In other words, we want to know when

$$\int_{-\infty}^{10-r} f(x)\,dx + \int_{10+r}^{\infty} f(x)\,dx < 0.01.$$

A bit of trial and error shows that with $r = 8$ the value is about 0.011, and with $r = 9$ it is about 0.004, so if the number of defective chips is 19 or more, or 1 or fewer, we should look for problems. If the number is high, we worry that the manufacturing process has a problem, or conceivably that the process that tests for defective chips is not working correctly and is flagging good chips as defective. If the number is too low, we suspect that the testing procedure is broken, and is not detecting defective chips. ♣

4.4.2. Two Random Variables

In Section 4.4.1, we have learned that the probability density function characterizes the distribution of a continuous random variable. Often, we are interested in several random variables that are related to each other, such as the cost of ski tickets and number of skiers at some local mountain resort. Using multi-variable calculus, we can generalize the ideas from Section 4.4.1 to two or more continuous random variables. In these notes, we will restrict ourselves to two continuous random variables X and Y.

4.4.2.1. Joint Probability Density and Joint Cumulative Distribution

A pair of continuous random variables is characterized by a so-called **joint probability density function**.

Definition 4.48: Joint Probability Density Function

Let f be an integrable function. Then f is the **joint probability density function** *of a pair of continuous random variables X and Y if f satisfies the following two properties:*

1. $f(x,y) \geq 0$ *for all x and for all y.*
2. $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$

Note:

1. Recall Fubini's Theorem 4.17, which says that the order of integration does not matter as long as the function which is integrated is continuous over the region of integration. Throughout this section, we simply write $dxdy$ rather than pointing out every time that we can choose between $dxdy$ and $dydx$.
2. At times, the joint probability density function of a pair of random variables X and Y is written as $f_{XY}(x,y)$.
3. The **marginal probability density functions** are given by
$$f_X(x,y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy \quad \text{and} \quad f_Y(x,y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx.$$
4. The pair of continuous random variables X and Y is **independent** if and only if the joint probability density function of X and Y factors into the product of their marginal probability density functions:
$$f_{XY}(x,y) = f_X(x) f_Y(y).$$

Example 4.49: Verifying a Joint Probability Density Function

Verify that
$$f(x,y) = \frac{2}{21}(4x+y)$$
is a joint probability density function on $[0,1]\times[0,3]$ for a pair of continuous random variables X and Y.

Solution. First, we observe that f is non-negative on $[0,1]\times[0,3]$.
Second, we evaluate

$$\int_0^3\int_0^1 \frac{2}{21}(4x+y)\,dx\,dy = \frac{2}{21}\int_0^3 \left[2x^2+xy\right]_0^1 dy = \frac{2}{21}\int_0^3 (2+y)\,dy$$
$$= \frac{2}{21}\left[2y+0.5y^2\right]_0^3 = \frac{2}{21}(10.5) = 1.$$

Hence, f is a joint probability density function on $[0,1]\times[0,3]$ for X and Y. ♣

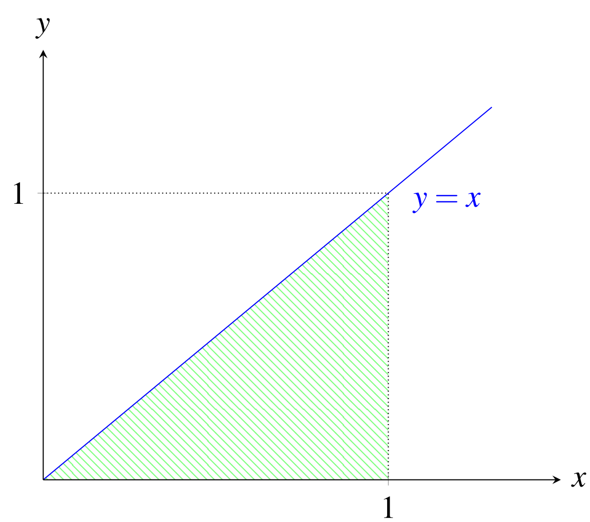
Example 4.50: Finding a Joint Probability Density Function

Find the constant c so that

$$f(x,y) = \begin{cases} cxy^2 & 0 \le y \le x \le 1 \\ 0 & \text{all other } x \text{ and } y \end{cases}$$

is a joint probability density function for the pair of continuous random variables X and Y.

Solution. We begin by graphing the region $0 \le y \le x \le 1$ in the x-y-plane:



Therefore we have that $y \le x \le 1$ with $0 \le y \le 1$ or $0 \le y \le x$ with $0 \le x \le 1$.

To find the constant c, we solve

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1.$$

Evaluating the double integral, we find

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dy\,dx$$
$$= \int_0^1\int_0^x cxy^2\,dy\,dx = \int_0^1 \left[\frac{xcy^3}{3}\right]_0^x dy$$
$$= \int_0^1 \frac{cx^4}{3}\,dx = \left.\frac{cx^5}{15}\right|_0^1 = \frac{c}{15}.$$

Hence, $c = 15$ and so

$$f(x,y) = \begin{cases} 15xy^2 & 0 \le y \le x \le 1 \\ 0 & \text{all other } x \end{cases}$$

is a joint probability function for X and Y. ♣

Example 4.51: Independent or Not

Let X and Y be a pair of continuous random variables with joint density function

$$f(x,y) = \begin{cases} 15xy^2 & 0 \le y \le x \le 1 \\ 0 & \text{all other } x \text{ and } y. \end{cases}$$

Are X and Y independent?

Solution. First, we compute

$$f_X(x) = \int_0^x 15xy^2\,dy = 3xy^3\big|_0^x = 3x^4,$$

and

$$f_Y(y) = \int_0^1 15xy^2\,dx = \frac{15}{2}x^2y^2\bigg|_0^1 = \frac{15}{2}y^2.$$

Since

$$f_{XY}(x,y) = 15xy^2 \neq f_X(x)f_Y(y) = 3x^4 \cdot \frac{15}{2}y^2,$$

we have that X and Y are not independent. ♣

As expected, the probability that the pair of continuous random variables X and Y lies in a region R is obtained by integrating its joint probability density function over R, which provides us with a **joint cumulative distribution function**.

Definition 4.52: Joint Cumulative Distribution Function

Suppose f is the probability density function of a pair of random variables X and Y. Then the **joint cumulative distribution function** *is*

$$F(x,y) = P(X \le x, Y \le y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f(v,w)\,dv\,dw.$$

Note:

1. Similar to the single random variable case, we have

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y)$$

for a pair of random variables X and Y. The derivative here is a mixed partial second derivative, which you should have encountered in any differential calculus course.

2. At times, the joint cumulative distribution function of a pair of random variables X and Y is written as $F_{XY}(x,y)$.

3. $F(-\infty,-\infty) = 0$ and $F(\infty,\infty) = 1$.

4. $F(x,y)$ is an increasing function in both x and y.

5. If the pair of random variables X and Y is **independent**, then
$$F_{XY}(x,y) = F_X(x)F_Y(y).$$

Example 4.53: Finding a Cumulative Distribution Function

Let X and Y be two independent uniform random variables on $[0,1]$. Find their joint cumulative distribution function $F_{XY}(x,y)$.

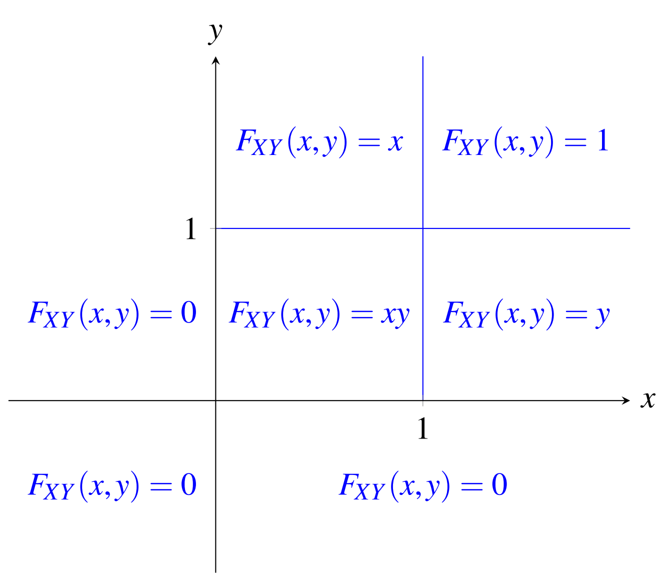
Solution. Since X and Y are uniform on $[0,1]$, we have that

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x \le 1 \\ 1 & x > 1 \end{cases} \quad \text{and} \quad F_Y(y) = \begin{cases} 0 & y < 0 \\ y & 0 \le y \le 1 \\ 1 & y > 1. \end{cases}$$

Since X and Y are independent, we obtain

$$F_{XY}(x,y) = \begin{cases} 0 & x < 0 \text{ or } y < 0 \\ xy & 0 \le x \le 1, 0 \le y \le 1 \\ x & 0 \le x \le 1, y > 1 \\ y & x > 1, 0 \le y \le 1 \\ 1 & x > 1, y > 1. \end{cases}$$

The graph below shows the values of the joint cumulative distribution function $F_{XY}(x,y)$ in the x-y-plane:



♣

Example 4.54: Finding a Cumulative Distribution Function

Let X and Y be a pair of continuous random variables with joint density function

$$f(x,y) = \begin{cases} 1.5x^2 + y & x \in [0,1],\ y \in [0,1] \\ 0 & \text{all other } x \text{ and } y. \end{cases}$$

Find the joint cumulative distribution function $F(x,y)$.

Solution. We first observe that $F(x,y) = 0$ for $x < 0$ or $y < 0$, and $F(x,y) = 1$ for $x \geq 1$ and $y \geq 1$.

We integrate the joint density function to find the cumulative distribution function for $x > 0$ and $y > 0$:

$$F(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} \left(1.5v^2 + w\right)\, dv\, dw = \int_{0}^{y} \int_{0}^{x} \left(1.5v^2 + w\right)\, dv\, dw.$$

When $0 \leq x \leq 1$ and $0 \leq y \leq 1$, then

$$\begin{aligned} F(x,y) &= \int_0^y \int_0^x \left(1.5v^2 + w\right)\, dv\, dw = \int_0^y \left[0.5v^3 + wv\right]_0^x\, dw \\ &= \int_0^y \left(0.5x^3 + wx\right)\, dw = \left[0.5x^3 w + 0.5w^2 x\right]_0^y \\ &= 0.5x^3 y + 0.5xy^2. \end{aligned}$$

When $0 \leq x \leq 1$ and $y \geq 1$, we use that $F(x,y)$ is continuous, then

$$F(x,y) = 0.5x^3 + 0.5x.$$

When $x \geq 1$ and $0 \leq y \leq 1$, we use that $F(x,y)$ is continuous, then

$$F(x,y) = 0.5y + 0.5y^2.$$

Hence,

$$f_{XY}(x,y) = \begin{cases} 0 & x < 0 \text{ or } y < 0 \\ 0.5x^3 y + 0.5xy^2 & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0.5x^3 + 0.5x & 0 \leq x \leq 1, y \geq 1 \\ 0.5y + 0.5y^2 & 0 \leq x \leq 1, y \leq 1 \\ 1 & x \geq 1, y \geq 1. \end{cases}$$

♣

Definition 4.55: Probability – Two Random Variables

Suppose f is the probability density function of a pair of random variables X and Y. Then the **probability** *that X and Y take values in the region $R = [a,b] \times [c,d] \subseteq \mathbb{R}^2$ is*

$$P(a \leq X \leq b, c \leq Y \leq d) = F(b,d) - F(a,d) - F(b,c) + F(a,c) = \int_c^d \int_a^b f(x,y)\, dx\, dy.$$

Example 4.56: Calculating a Probability

Given the probability density function

$$f(x,y) = \begin{cases} 1 & (x,y) \in [0,1] \times [0,1] \\ 0 & \text{all other } (x,y) \end{cases}$$

of a pair of continuous random variables X and Y, calculate the following probabilities:

(a) $P(X = 1/2, Y = 1/2)$ *(b)* $P(X \leq 1/2, Y \leq 1/2)$

Solution.

(a) The probability is computed as

$$P(X = 1/2, Y = 1/2) = \int_{1/2}^{1/2} \int_{1/2}^{1/2} dx\,dy = \int_{1/2}^{1/2} x\Big|_{1/2}^{1/2} dy = \int_{1/2}^{1/2} 0\,dy = 0.$$

(b) The probability is computed as

$$\begin{aligned} P(X \leq 1/2, Y \leq 1/2) &= P((X,Y) \in (-\infty, 1/2] \times (-\infty, 1/2]) \\ &= \int_{-\infty}^{1/2} \int_{-\infty}^{1/2} f(x,y)\,dx\,dy = \int_{0}^{1/2} \int_{0}^{1/2} dx\,dy \\ &= \int_{0}^{1/2} x\Big|_0^{1/2} dy = \int_0^{1/2} \frac{1}{2}\,dy = \frac{1}{2} y\Big|_0^{1/2} = \frac{1}{4}. \end{aligned}$$

♣

4.4.2.2. Expected Value, Variance and Covariance

Analogous to the single random variable case, we can compute the expected value, variance and standard deviation for a pair of continuous random variables.

Definition 4.57: Expected Values for a Pair of Continuous Random Variables

Suppose X and Y are a pair of continuous random variables with probability density function $f(x,y)$.

Then the **expected value** *of X is*

$$E(X) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x,y)\,dx\,dy$$

and the **expected value** *of Y is*

$$E(Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x,y)\,dx\,dy$$

provided the integrals converge.

Note: Equivalently, we can use the marginal probability density functions to compute the expected values of X and Y respectively:

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx \quad \text{and} \quad E(Y) = \int_{-\infty}^{\infty} y f_Y(y)\, dy.$$

Definition 4.58: Variance of a Pair for Continuous Random Variables

Suppose X and Y are a pair of continuous random variables with probability density function $f(x,y)$.

Then the **variance** *of X is*

$$V(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 f(x,y)\, dx\, dy - \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f(x,y)\, dx\, dy\right)^2$$

and the **variance** *of Y is*

$$V(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y^2 f(x,y)\, dx\, dy - \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(x,y)\, dx\, dy\right)^2$$

provided the integrals converge.

Note: Since the calculation of the variance is based on the mean, we can write

$$V(X) = E(X^2) - E(X)^2 \quad \text{and} \quad V(Y) = E(Y^2) - E(Y)^2.$$

Definition 4.59: Standard Deviation for a Pair of Continuous Random Variables

Suppose X and Y are a pair of continuous random variables with probability density function $f(x,y)$.

Then the **standard deviation** *of X and the standard deviation of Y are*

$$\sigma(X) = \sqrt{V(X)} \quad \text{and} \quad \sigma(Y) = \sqrt{V(Y)}$$

respectively, where $V(X)$ is the variance of X and $V(Y)$ is the variance of Y.

Example 4.60: Calculating Expected Values and Variance

Let X and Y be a pair of continuous random variables with probability density function

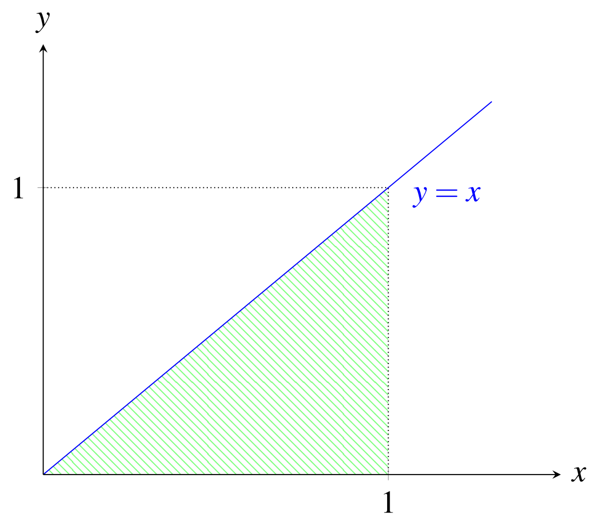
$$f(x,y) = \begin{cases} 15xy^2 & 0 \le y \le x \le 1 \\ 0 & \text{all other } x \text{ and } y. \end{cases}$$

(a) What is the expected value of X?

(b) What is the expected value of Y?

(c) Compute $V(X)$ and $V(Y)$.

Solution. Recall from Example 4.50 that the region $0 \le y \le x \le 1$ in the x-y-plane is graphed as follows:



Therefore we have that $y \le x \le 1$ with $0 \le y \le 1$ or $0 \le y \le x$ with $0 \le x \le 1$. Recall that the order of integration does not matter for continuous f. In our work below, we simply want to show both orders of integration.

(a) The expected value of X is

$$\begin{aligned} E(X) &= \int_0^1 \int_y^1 x\left(15xy^2\right)\,dx\,dy = \int_0^1 \left[5x^3y^2\right]_y^1 dy \\ &= 5\int_0^1 \left(y^2 - y^5\right) dy = 5\left[\frac{y^3}{3} - \frac{y^6}{6}\right]_0^1 = \frac{5}{6}. \end{aligned}$$

(b) The expected value of Y is

$$\begin{aligned} E(Y) &= \int_0^1 \int_0^x y\left(15xy^2\right)\,dy\,dx = \int_0^1 \left[\frac{15xy^4}{4}\right]_0^x dx \\ &= \int_0^1 \frac{15x^5}{4}\,dx = \left.\frac{5x^6}{8}\right|_0^1 = \frac{5}{8}. \end{aligned}$$

(c) Using $V(X) = E(X^2) - E(X)^2$, we obtain

$$\begin{aligned} E(X^2) &= \int_0^1 \int_y^1 x^2\left(15xy^2\right)\,dx\,dy = \int_0^1 \left[\frac{15}{4}x^4y^2\right]_y^1 dy \\ &= \frac{15}{4}\int_0^1 \left(y^2 - y^6\right) dy = \frac{15}{4}\left[\frac{y^3}{3} - \frac{y^7}{7}\right]_0^1 = \frac{15}{4}\cdot\frac{4}{21} = \frac{5}{7}. \end{aligned}$$

Using the result from (a), we have

$$V(X) = E(X^2) - E(X)^2 = \frac{5}{7} - \left(\frac{5}{6}\right)^2 = \frac{5}{252}.$$

Similarly, we compute

$$E(Y^2) = \int_0^1 \int_0^x y^2 \left(15xy^2\right) dy\,dx = \int_0^1 \left[3xy^5\right]_0^x dx$$
$$= \int_0^1 3x^6\,dx = \left.\frac{3x^7}{7}\right|_0^1 = \frac{3}{7},$$

and using the result from (b), we have

$$V(Y) = E(Y^2) - E(Y)^2 = \frac{3}{7} - \left(\frac{5}{8}\right)^2 = \frac{17}{448}.$$

♣

Aside from the expected value, variance and standard deviation, we may also be interested in how two continuous random variables are related. To measure such a relationship, we define the **covariance**.

Definition 4.61: Covariance for a Pair of Continuous Random Variables

Suppose X and Y are a pair of continuous random variables with probability density function $f(x,y)$. Then the **covariance** *of X and Y is*

$$Cov(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\,dx\,dy - \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)\,dx\,dy\right)\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)\,dx\,dy\right)$$

provided the integrals converge.

Note:

1. The sign of the covariance of two random variables X and Y tells us the direction of the linear relationship between them:
   - If the covariance is *positive*, we say X and $Y$are positively correlated. This means that large values of X tend to happen with large values of Y, and similarly small values of X tend to happen with small values of Y.
   - If the covariance is *negative*, we say X and Y are negatively correlated. This means that small values of X tend to happen with large values of Y, and vice versa.
2. Since the calculation of the covariance is based on the mean, we can write
$$Cov(X,Y) = E(XY) - E(X)E(Y).$$
3. $Cov(X,X) = V(X)$.
4. If the random variables X and Y are independent, then $Cov(X,Y) = 0$.

Example 4.62: Calculating Covariance

Let X and Y be a pair of continuous random variables with probability density function

$$f(x,y) = \begin{cases} 15xy^2 & 0 \le y \le x \le 1 \\ 0 & \text{all other } x \text{ and } y. \end{cases}$$

Compute the covariance of X and Y, and interpret your result.

Solution. We begin by computing

$$\int_0^1 \int_y^1 xy(15xy^2)\,dx\,dy = \int_0^1 \left[5x^3y^3\right]_y^1 dy = 5\int_0^1 \left(y^3 - y^6\right) dy$$
$$= 5\left[\frac{y^4}{4} - \frac{y^7}{7}\right]_0^1 = 5\left(\frac{1}{4} - \frac{1}{7}\right) = \frac{15}{28}.$$

Using the results from Example 4.60, we find the covariance of X and Y:

$$Cov(X,Y) = \int_0^1 \int_y^1 xy(15xy^2)\,dx\,dy - E(X)E(Y)$$
$$= \frac{15}{28} - \frac{5}{6}\cdot\frac{5}{8} = \frac{15}{28} - \frac{25}{48} = \frac{5}{336} \approx 0.0149.$$

The covariance is slightly positive. Hence, large values of X tend to occur more often with large values of Y. ♣

Exercises for Section 4.4

Exercise 4.4.1 *Verify that f is a probability density function on the given interval.*

(a) $f(x) = \dfrac{10}{3x^2}, \quad x \in [2,5]$

(b) $f(x) = 6(\sqrt{x} - x), \quad x \in [0,1]$

(c) $f(x,y) = \dfrac{1}{3}, \quad (x,y) \in [1,2] \times [3,6]$

(d) $f(x,y) = \dfrac{1}{4}xy, \quad (x,y) \in [0,1] \times [0,4]$

Exercise 4.4.2 *Construct a probability density function f from the given function g.*

(a) $g(x) = 3 - x, \quad x \in [0,3]$

(b) $g(x) = \dfrac{1}{x^5}, \quad x \in [1,\infty)$

(c) $g(x,y) = xy^2, \quad (x,y) \in [0,1] \times [0,1]$

(d) $g(x,y) = x^2 e^{-y}, \quad (x,y) \in [1,2] \times [1,\infty)$

Exercise 4.4.3 *Calculate the following cumulative distributions.*

(a) $f(x) = \dfrac{1}{2}e^{-x/2}, \quad x \in [0,\infty)$

i. $P(x = 1)$

ii. $P(3 \leq x \leq 6)$

iii. $P(x \leq 50)$

iv. $P(x \geq 6)$

(b) $f(x) = \frac{3}{14}\sqrt{x}, \quad x \in [1,4]$

i. $P(2 \leq x \leq 4)$

ii. $P(1 \leq x \leq 3)$

iii. $P(x \leq 2)$

iv. $P(x \geq 2)$

(c) $f(x,y) = \frac{1}{3}xy, \quad (x,y) \in [0,2] \times [1,2]$

i. $P(0 \leq x \leq 1, 1 \leq y \leq 2)$

ii. $P(1 \leq x \leq 2, y = 1)$

(d) $f(x,y) = \frac{1}{16}(2-x)y, \quad (x,y) \in [0,2] \times [0,4]$

i. $P(0 \leq x \leq 1, 0 \leq y \leq 1)$

ii. $P(x \geq 1, y \geq 3)$

Exercise 4.4.4 *Let*

$$f(x) = \begin{cases} \frac{1}{x^2} & x \geq 1 \\ 0 & x < 1 \end{cases}$$

Show that f is a probability density function, and that the distribution has no mean.

Exercise 4.4.5 *Let*

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

Show that $\int_{-\infty}^{\infty} f(x)\,dx = 1$. *Is f a probability density function? Justify your answer.*

Exercise 4.4.6 *A sawmill wants to assess the performance of a debarker. They determine that length of time between machine failures is exponentially distributed with probability density function*

$$f(t) = 0.005e^{-0.005t},$$

where t is measured in hours.

(a) Determine the probability that the debarker breaks down between $t = 200$ and $t = 600$ hours.

(b) Determine the probability that the machine breaks down after $t = 1000$ hours.

Exercise 4.4.7 *For each of the given probability density functions $f(x)$, determine (i) the mean, (ii) the variance, and (iii) the standard deviation.*

(a) $f(x) = \frac{3}{215}x^2,\ x \in [1,6]$

(b) $f(x) = \frac{1}{36}(x-1)(7-x),\ x \in [1,7]$

(c) $f(x) = \frac{24}{x^4},\ x \in [2,\infty)$

(d) $f(x) = \frac{1}{25}e^{-x/5},\ x \in [0,\infty)$

Exercise 4.4.8 *Suppose the probability density function which describes the number of leaves on a certain plant is given by*

$$f_\ell = \frac{1}{20}\left(5\ell - \ell^2\right), \quad \ell = 0,1,...,5.$$

(a) Determine the probability that a plant has exactly one leaf.

(b) Determine the probability that a plant has less than 3 leaves.

Exercise 4.4.9 *A city planner wishes to improve traffic on a busy bridge. He determines that the length of time it takes a car to cross the bridge is a continuous random variable with probability density function*

$$f(t) = \frac{4}{(4+t^2)^{3/2}}, \quad t \in [0,\infty),$$

where t is measured in minutes. How long is a randomly chosen car expected to take to cross the bridge?

Exercise 4.4.10 *The amount of chicken (in kg) demanded weekly at a popular restaurant is a continuous random variable with probability distribution function*

$$f(x) = \frac{3}{500}x(10-x), \quad x \in [0,5].$$

What is the expected weekly demand for chicken?

Exercise 4.4.11 *Let X and Y be a pair of continuous random variables with the given joint probability density function. Are X and Y independent?*

(a) $f_{XY}(x,y) = \begin{cases} 6xy & 0 \le x \le 1, 0 \le y \le \sqrt{x} \\ 0 & \text{all other } x \text{ and } y. \end{cases}$

(b) $f_{XY}(x,y) = \begin{cases} 6e^{-(2x+3y)} & x,y \ge 0 \\ 0 & \text{all other } x \text{ and } y. \end{cases}$

(c) $f_{XY}(x,y) = \begin{cases} 10x^2y & 0 \le y \le x \le 1 \\ 0 & \text{all other } x \text{ and } y. \end{cases}$

(d) $f_{XY}(x,y) = \begin{cases} 2 & 0 \le x \le y \le 1 \\ 0 & \text{all other } x \text{ and } y. \end{cases}$

Exercise 4.4.12 *Given the following probability density function $f(x,y)$ for a pair of continuous random variables X and Y, compute (i) the expected values $E(X)$ and $E(Y)$, (ii) the variance $V(X)$ and $V(Y)$, and (iii) the covariance $Cov(X,Y)$.*

(a) $f_{XY}(x,y) = \begin{cases} 6xy & 0 \le x \le 1, 0 \le y \le \sqrt{x} \\ 0 & \textit{all other } x \textit{ and } y. \end{cases}$

(b) $f_{XY}(x,y) = \begin{cases} 6e^{-(2x+3y)} & x, y \ge 0 \\ 0 & \textit{all other } x \textit{ and } y. \end{cases}$

(c) $f_{XY}(x,y) = \begin{cases} 10x^2y & 0 \le y \le x \le 1 \\ 0 & \textit{all other } x \textit{ and } y. \end{cases}$

(d) $f_{XY}(x,y) = \begin{cases} 2 & 0 \le x \le y \le 1 \\ 0 & \textit{all other } x \textit{ and } y. \end{cases}$

5. Differential Equations

Many physical phenomena can be modeled using the language of calculus. For example, observational evidence suggests that the temperature of a cup of tea (or some other liquid) in a room of constant temperature will cool over time at a rate proportional to the difference between the room temperature and the temperature of the tea.

In symbols, if t is the time, M is the room temperature, and $f(t)$ is the temperature of the tea at time t then $f'(t) = k(M - f(t))$ where $k > 0$ is a constant which will depend on the kind of tea (or more generally the kind of liquid) but not on the room temperature or the temperature of the tea. This is **Newton's law of cooling** and the equation that we just wrote down is an example of a **differential equation**. Ideally we would like to solve this equation, namely, find the function $f(t)$ that describes the temperature over time, though this often turns out to be impossible, in which case various approximation techniques must be used. The use and solution of differential equations is an important field of mathematics, because differential equations help us to *predict* future behaviour based on how current values are related and how they change with respect to each other (perhaps over time). Here we see how to solve some simple but useful types of differential equation.

Informally, a differential equation is an equation in which one or more of the derivatives of some function appears. Typically, a scientific theory will produce a differential equation (or a system of differential equations) that describes or governs some physical process, but the theory will not produce the desired function or functions directly.

Definition 5.1: Differential Equation

A **differential equation** *is a mathematical equation for an unknown function of one (or several) variables that relates the function to its derivatives.*

A **solution** *to a differential equation is a function that satisfies the differential equation.*

Note: The term *Differential Equation* is often abbreviated with **DE**, and so DEs stands for differential equations.

The following are examples of differential equations:

1. $y' = e^x \sec y$
2. $y' - e^x y + 3 = 0$
3. $y' - e^x y = 0$
4. $3y'' - 2y' = 7$
5. $4\dfrac{d^5 y}{dx^5} + \cos x \dfrac{dy}{dx} = 0$

Clearly, there are many different characteristics of a differential equation. The characteristics that are used throughout the notes are introduced below. However, there are additional ways to classify differential equations, which we leave to the interested reader, who pursues this field of study.

5.1 Classifying Differential Equations

Definition 5.2: Order of a DE

The **order** *of a differential equation is the order of the largest derivative that appears in the equation.*

Let's come back to our list of examples and state the order of each differential equation:

1. $y' = e^x \sec y$ has order 1
2. $y' - e^x y + 3 = 0$ has order 1
3. $y' - e^x y = 0$ has order 1
4. $3y'' - 2y' = 7$ has order 2
5. $4\dfrac{d^5y}{dx^5} + \cos x \dfrac{dy}{dx} = 0$ has order 5

Definition 5.3: Linearity of a DE

A **linear** *differential equation can be written in the form*

$$F_n(x)\frac{d^ny}{dx^n} + F_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + F_2(x)\frac{d^2y}{dx^2} + F_1(x)\frac{dy}{dx} + F_0(x)y = G(x)$$

where $F_i(x)$ and $G(x)$ are functions of x. Otherwise, we say that the differential equation is **non-linear**.

Note: As an aside, if the leading coefficient $F_n(x)$ is non-zero, then the equation is said to be of n-th order.

Let's come back to our list of differential equations and add whether it is linear or not:

1. $y' = e^x \sec y$ has order 1, is non-linear
2. $y' - e^x y + 3 = 0$ has order 1, is linear
3. $y' - e^x y = 0$ has order 1, is linear
4. $3y'' - 2y' = 7$ has order 2, is linear

5. $4\dfrac{d^5y}{dx^5} + \cos x \dfrac{dy}{dx} = 0$ has order 5, is linear

Definition 5.4: Homogeneity of a Linear DE

Given a linear differential equation

$$F_n(x)\frac{d^ny}{dx^n} + F_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + F_2(x)\frac{d^2y}{dx^2} + F_1(x)\frac{dy}{dx} + F_0(x)y = G(x)$$

where $F_i(x)$ and $G(x)$ are functions of x, the differential equation is said to be **homogeneous** *if $G(x) = 0$ and* **non-homogeneous** *otherwise.*

Note: One implication of this definition is that $y = 0$ is a constant solution to a linear homogeneous differential equation, but not for the non-homogeneous case.

Let's come back to all linear differential equations on our list and label each as homogeneous or non-homogeneous:

2. $y' - e^x y + 3 = 0$ has order 1, is linear, is non-homogeneous
3. $y' - e^x y = 0$ has order 1, is linear, is homogeneous
4. $3y'' - 2y' = 7$ has order 2, is linear, is non-homogeneous
5. $4\dfrac{d^5y}{dx^5} + \cos x \dfrac{dy}{dx} = 0$ has order 5, is linear, is homogeneous

Definition 5.5: Linear DE with Constant Coefficients

Given a linear differential equation

$$F_n(x)\frac{d^ny}{dx^n} + F_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + F_2(x)\frac{d^2y}{dx^2} + F_1(x)\frac{dy}{dx} + F_0(x)y = G(x)$$

where $G(x)$ is a function of x, the differential equation is said to have **constant coefficients** *if $F_i(x)$ are constants for all i.*

As examples, we identify all linear differential equations on our list that have constant coefficients:

2. $y' - e^x y + 3 = 0$ has order 1, is linear, is non-homogeneous, does not have constant coefficients
3. $y' - e^x y = 0$ has order 1, is linear, is homogeneous, does not have constant coefficients
4. $3y'' - 2y' = 7$ has order 2, is linear, is non-homogeneous, has constant coefficients
5. $4\dfrac{d^5y}{dx^5} + \cos x \dfrac{dy}{dx} = 0$ has order 5, is linear, is homogeneous, does not have constant coefficients

Example 5.6: Newton's Law of Cooling

The equation from Newton's law of cooling,

$$\frac{dy}{dt} = k(M - y)$$

is a first order linear non-homogeneous differential equation with constant coefficients, where t is time, k is the constant of proportionality, and M is the ambient temperature.

Exercises for Section 5.1

Exercise 5.1.1 *Identify the order and linearity of each differential equation below.*

(a) $5y''' + 3y' - 4\sin(y) = \cos(x)$

(b) $4\dfrac{d^3y}{dx^3} + 2\dfrac{dy}{dx} = e^x y$

(c) $e^x \dfrac{dy}{dx} + e^{x+y} = e$

(d) $-\dfrac{d^4y}{dx^4} + x\dfrac{d^3y}{dx^3} - x^2\dfrac{d^2y}{dx^2} + x^3\dfrac{dy}{dx} - x^4y^5 = 0$

(e) $f_2(x)y'' + f_1(x)y' + f_0(x)y - 1 = 0$, *where* $f_i(x)$ *are non-constant functions of* x

(f) $\tan(xy)y' = \sec(xy)$

Exercise 5.1.2 *Identify the homogeneity of each linear differential equation below.*

(a) $5y''' + 3y' - 4y = \cos(x)$

(b) $4\dfrac{d^3y}{dx^3} + 2\dfrac{dy}{dx} = e^x y$

(c) $e^x \dfrac{dy}{dx} + e^x y = e$

(d) $-\dfrac{d^4y}{dx^4} + x\dfrac{d^3y}{dx^3} - x^2\dfrac{d^2y}{dx^2} + x^3\dfrac{dy}{dx} - x^4y = 0$

(e) $f_2(x)y'' + f_1(x)y' + f_0(x)y - 1 = 0$, *where* $f_i(x)$ *are non-constant functions of* x

(f) $\tan(x)y' = \sec(x)y$

Exercise 5.1.3 *State whether the coefficients of each linear differential equation below are constant or not.*

(a) $5y''' + 3y' - 4y = \cos(x)$

(b) $4\dfrac{d^3y}{dx^3} + 2\dfrac{dy}{dx} = e^x y$

(c) $e^x\dfrac{dy}{dx} + e^x y = e$

(d) $-\dfrac{d^4y}{dx^4} + x\dfrac{d^3y}{dx^3} - x^2\dfrac{d^2y}{dx^2} + x^3\dfrac{dy}{dx} - x^4 y = 0$

(e) $f_2(x)y'' + f_1(x)y' + f_0(x)y - 1 = 0$, *where* $f_i(x)$ *are non-constant functions of* x

(f) $\tan(x)y' = \sec(x)y$

5.2 First Order Differential Equations

In many fields such as physics, biology or business, a relationship is often known or assumed between some unknown quantity and its rate of change, which does not involve any higher derivatives. It is therefore of interest to study first order differential equations in particular.

Definition 5.7: First Order DE

A **first order differential equation** *is an equation of the form* $F(t, y, y') = 0$. *A solution of a first order differential equation is a function* $f(t)$ *that makes* $F(t, f(t), f'(t)) = 0$ *for every value of* t.

Here, F is a function of three variables which we label t, y, and y'. It is understood that y' will explicitly appear in the equation although t and y need not. The variable y itself is dependent on t, hence it is understood that y' must be the derivative of y with respect to t. Since only the first derivative of y appears, but no higher order derivative, this is a *first order* differential equation.

Throughout the notes, we use the independent variable t as many applications are based on the independent variable representing time. If no meaning is attributed to the independent variable, we may want to write a first order differential equation in the usual manner as

$$F(x, y, y') = 0.$$

Example 5.8: Simple First Order Differential Equation

$y' = t^2 + 1$ *is a first order linear differential equation;* $F(t, y, y') = y' - t^2 - 1$. *Show that all solutions to this equation are of the form* $y = t^3/3 + t + C$.

Solution. We first note that $y = t^3/3 + t + C$ is a solution to the differential equation, since

$$\frac{d}{dt}\left(\frac{t^3}{3} + t + C\right) = t^2 + 1,$$

for all $C \in \mathbb{R}$.

We additionally need to show that there are no other solutions. To do so, we integrate the differential equation:

$$y(t) = \int \left(t^2 + 1\right)\, dt = \frac{t^3}{3} + t + C,$$

for some $C \in R$.

Thus, all solutions to the differential equation are of the form $y = t^3/3 + t + C$. ♣

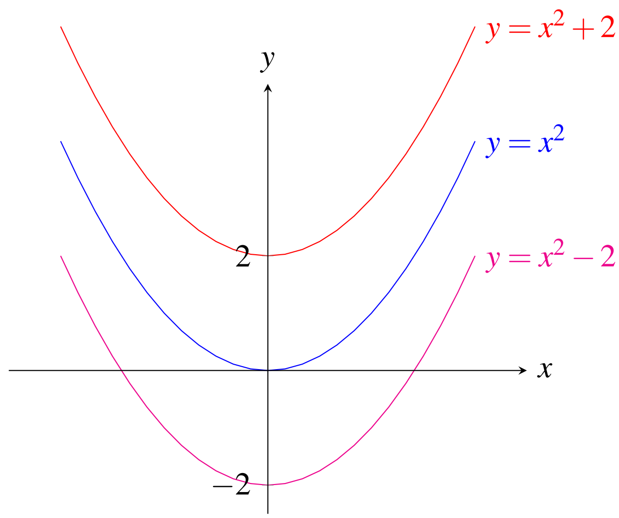
Example 5.9: Graphical Solution to First Order Differential Equation

Sketch various solutions to the differential equation $\dfrac{dy}{dx} = 2x$.

Solution. We integrate both sides of the differential equation to find

$$y = \int 2x\, dx = x^2 + C$$

for any constant $C \in \mathbb{R}$. This family of solutions are parabolas which are translated vertically, as shown in the graph below taking $C = -2, 0, 2$.



♣

5.2.1. Initial Value Problems

Definition 5.10: Initial Conditions

Initial condition(s) *are a set of points that the solution (or its derivatives) must satisfy.*

Example 5.11: Initial Conditions

For a differential equation involving a function $f(t)$, initial conditions are of the form:

$$f(t_0) = f_0, \qquad f'(t_0) = f_1, \qquad f''(t_0) = f_2, \qquad \ldots etc.$$

Definition 5.12: Initial Value Problem

An ***initial value problem*** *(IVP) is a differential equation along with a set of initial conditions.*

Example 5.13: First Order Initial Value Problem

Solve the initial value problem:

$$\frac{dy}{dx} = 2x, \quad y(0) = 2.$$

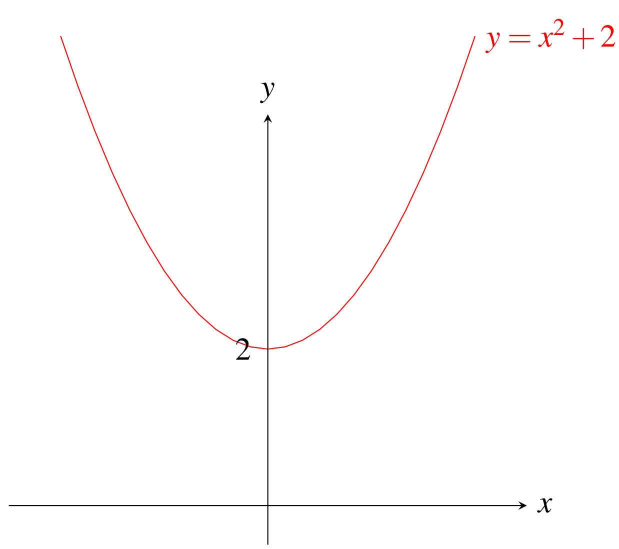
Solution. We had found in Example 5.9 that the solutions to the differential equation were the parabolas

$$y(x) = x^2 + C.$$

So we use the initial condition to determine the constant C:

$$y(0) = (0)^2 + C = 2 \implies C = 2.$$

Therefore, the solution to the initial value problem is $y = x^2 + 2$, as shown in the graph below.



♣

Example 5.14: Simple Initial Value Problem

Verify that the initial value problem $y' = t^2 + 1$, $y(1) = 4$ *has solution* $f(t) = t^3/3 + t + 8/3$.

Solution. Observe that $f'(t) = t^2 + 1$ and $f(1) = 1^3/2 + 1 + 8/3 = 4$ as required. ♣

The general first order equation is too general, so we can't describe methods that will work on them all, or even a large portion of them. We can make progress with specific kinds of first order differential equations. For example, much can be said about equations of the form $y' = \phi(t,y)$ where ϕ is a function of the two variables t and y. Under reasonable conditions on ϕ, such an equation has a solution and the corresponding initial value problem has a unique solution. However, in general, these equations can be very difficult or impossible to solve explicitly.

Example 5.15: IVP for Newton's Law of Cooling

Consider this specific example of an initial value problem for Newton's law of cooling:

$$y' = 2(25 - y), \quad y(0) = 40.$$

Discuss the solutions for this initial value problem.

Solution. We first note the zero of the equation: If $y(t_0) = 25$, the right hand side of the differential equation is zero, and so the constant function $y(t) = 25$ is a solution to the differential equation. It is not a solution to the initial value problem, since $y(0) \neq 25$. (The physical interpretation of this constant solution is that if a liquid is at the same temperature as its surroundings, then the liquid will stay at that temperature.) So long as y is not 25, we can rewrite the differential equation as

$$\begin{aligned} \frac{dy}{dt}\frac{1}{25-y} &= 2 \\ \frac{1}{25-y}\,dy &= 2\,dt, \end{aligned}$$

so

$$\int \frac{1}{25-y}\,dy = \int 2\,dt,$$

that is, the two antiderivatives must be the same except for a constant difference. We can calculate these antiderivatives and rearrange the results:

$$\begin{aligned} \int \frac{1}{25-y}\,dy &= \int 2\,dt \\ (-1)\ln|25-y| &= 2t + C_0 \\ \ln|25-y| &= -2t - C_0 = -2t + C \\ |25-y| &= e^{-2t+C} = e^{-2t}e^{C} \\ y - 25 &= \pm e^{C}e^{-2t} \\ y &= 25 \pm e^{C}e^{-2t} = 25 + Ae^{-2t}. \end{aligned}$$

Here $A = \pm e^C = \pm e^{-C_0}$ is some non-zero constant. Since we want $y(0) = 40$, we substitute and solve for A:

$$40 = 25 + Ae^0 \implies A = 15.$$

Therefore, $y = 25 + 15e^{-2t}$ is a solution to the initial value problem. Note that y is never 25, so this makes sense for all values of t. However, if we allow $A = 0$ we get the solution $y = 25$ to the differential equation, which would be the solution to the initial value problem if we were to require $y(0) = 25$. Thus, $y = 25 + Ae^{-2t}$ describes all solutions to the differential equation $y' = 2(25 - y)$, and all solutions to the associated initial value problems. ♣

5.2.2. Separable Equations

Why could we solve Example 5.15 from the previous section? Our solution depended on rewriting the equation so that all instances of y were on one side of the equation and all instances of t were on the other. Of course, in this case the only t was originally hidden, since we didn't write dy/dt in the original equation. This is not required, however. This idea of being able to separate the independent and dependent variables in a first order differential equation leads to a classification of first order differential equations into **separable** and **non-separable** equations as follows.

Definition 5.16: Separable DE

A first order differential equation is **separable** *if it can be written in the form*

$$\frac{dy}{dt} = f(t)g(y).$$

Let's come back to all first order differential equations on our list from the previous section and decide which ones are separable or not:

1. $y' = e^x \sec y$ has order 1, is non-linear, is separable
2. $y' - e^x y + 3 = 0$ has order 1, is linear, is not separable
3. $y' - e^x y = 0$ has order 1, is linear, is separable

As in the examples, we can attempt to solve a separable equation by converting to the form

$$\int \frac{1}{g(y)}\, dy = \int f(t)\, dt.$$

This technique is called **separation of variables**. The simplest (in principle) sort of separable equation is one in which $g(y) = 1$, in which case we attempt to solve

$$\int 1\, dy = \int f(t)\, dt.$$

We can do this if we can find an antiderivative of $f(t)$.

As we have seen so far, a differential equation typically has an infinite number of solutions. Such a solution is called a **general solution**. A corresponding initial value problem will give rise to just one solution. Such a solution in which there are no unknown constants remaining is called a **specific solution**.

The general approach to separable equations is as follows: Suppose we wish to solve $y' = f(t)g(y)$ where f and g are continuous functions. If $g(a) = 0$ for some a then $y(t) = a$ is a constant solution of the equation, since in this case $y' = 0 = f(t)g(a)$. For example, $y' = y^2 - 1$ has constant solutions $y(t) = 1$ and $y(t) = -1$.

To find the non-constant solutions, we note that the function $1/g(y)$ is continuous where $g \neq 0$, so $1/g$ has an antiderivative G. Let F be an antiderivative of f. Now we write

$$G(y) = \int \frac{1}{g(y)}\,dy = \int f(t)\,dt = F(t) + C,$$

so $G(y) = F(t) + C$. Now we solve this equation for y.

Of course, there are a few places this ideal description could go wrong: Finding the antiderivatives G and F, and solving the final equation for y. The upshot is that the solutions to the original differential equation are the constant solutions, if any, and all functions y that satisfy $G(y) = F(t) + C$.

Guideline for Separation of Variables

Given the differential equation

$$\frac{dy}{dt} = f(t)g(y),$$

follow these steps to find the non-constant solutions.

1. Separate the variables:
$$\frac{dy}{g(y)} = f(t)\,dt$$
2. Apply the integration operator:
$$\int \frac{dy}{g(y)} = \int f(t)\,dt$$
3. If an antiderivative exists for f and for $1/g$, and we can solve for y, then
$$G(y) = F(t) + C$$
for some constant C.

Example 5.17: Solving a Separable Differential Equation I

Solve the differential equation $y' = 2t(25 - y)$.

Solution. This is almost identical to the previous example. As before, $y(t) = 25$ is a solution. If $y \neq 25$,

$$\int \frac{1}{25 - y}\,dy \quad = \quad \int 2t\,dt$$

$$\begin{aligned}(-1)\ln|25-y| &= t^2+C_0\\ \ln|25-y| &= -t^2-C_0=-t^2+C\\ |25-y| &= e^{-t^2+C}=e^{-t^2}e^C\\ y-25 &= \pm e^Ce^{-t^2}\\ y &= 25\pm e^Ce^{-t^2}=25+Ae^{-t^2}.\end{aligned}$$

As before, all solutions are represented by $y=25+Ae^{-t^2}$, allowing A to be zero. ♣

Example 5.18: Solving a Seperable Differential Equation II

Find the solutions to the differential equation

$$\sec(t)\frac{dy}{dt}-e^{y+\sin(t)}=0.$$

Solution. We begin by separating the variables and get

$$\begin{aligned}\sec(t)\frac{dy}{dt} &= e^{y+\sin(t)}\\ \sec(t)\frac{dy}{dt} &= e^ye^{\sin(t)}\\ e^{-y}\,dy &= \frac{e^{\sin(t)}}{\sec(t)}\,dt=\cos(t)e^{\sin(t)}\,dt\end{aligned}$$

Now integrate both sides to obtain

$$\begin{aligned}\int e^{-y}\,dy &= \int\cos(t)e^{\sin(t)}\,dt\\ -e^{-y} &= e^{\sin(t)}+C\\ y &= -\ln\left(D-e^{\sin(t)}\right)\end{aligned}$$

For convenience, we left out the absolute value in the argument of the logarithm. As in the previous examples, care must be taken to ensure that the argument of the logarithm is positive for a given value of D.

Therefore, the solutions to the differential equation are given by

$$y=-\ln\left(D-e^{\sin(t)}\right),$$

for some constant D. ♣

5.2.3. Simple Growth and Decay Model

Example 5.19: Rate of Change Proportional to Size

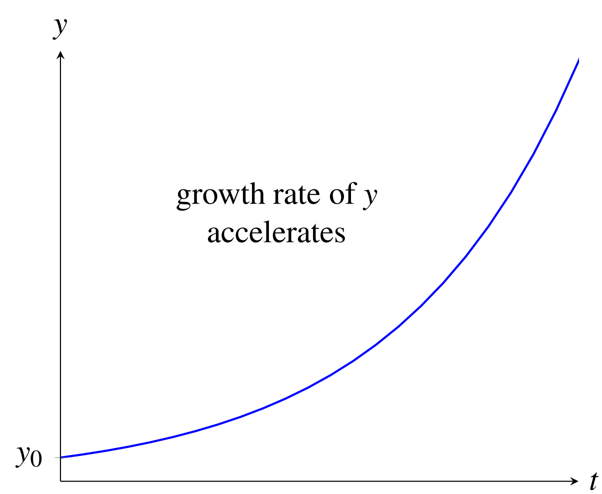
Find the solutions to the differential equation $y' = ky$, which models a quantity y that grows or decays proportionally to its size depending on whether k is positive or negative.

Solution. The constant solution is $y(t) = 0$; of course this will not be the solution to any interesting initial value problem. For the non-constant solutions, we proceed much as before:

$$\begin{aligned}
\int \frac{1}{y}\,dy &= \int k\,dt \\
\ln|y| &= kt + C \\
|y| &= e^{kt}e^{C} \\
y &= \pm e^{C}e^{kt} \\
y &= Ae^{kt}.
\end{aligned}$$

Again, if we allow $A = 0$ this includes the constant solution, and we can simply say that $y = Ae^{kt}$ is the general solution. With an initial value we can easily solve for A to get the solution of the initial value problem. In particular, if the initial value is given for time $t = 0$, $y(0) = y_0$, then $A = y_0$ and the solution is $y = y_0e^{kt}$. ♣

The constant k in the above differential equation is referred to as the **growth rate constant**. Furthermore, this type of differential equation is known as a simple model for growth and decay of some quantity, since it only considers that the growth rate is proportional to the size of the quantity itself without any other factors influencing y. The graph below shows the typical J-shape of such a solution for some y_0.



Simple Growth and Decay Model

The differential equation
$$\frac{dy}{dt} = ky$$
with growth rate constant k models **simple growth and decay** of a quantity y at time t with solution
$$y = y_0 e^{kt},$$
where y_0 is the initial value at time $t = 0$.

When $k > 0$, this describes certain simple cases of population growth: It says that the change in the population y is proportional to the population. The underlying assumption is that each organism in the current population reproduces at a fixed rate, so the larger the population the more new organisms are produced. While this is too simple to model most real populations, it is useful in some cases over a limited time.

When $k < 0$, the differential equation describes a quantity that decreases in proportion to the current value. This can be used to model radioactive decay.

Note: The simple growth and decay model is unrestricted because the quantity y grows without bound as $t \to \infty$ if $k > 0$. In the decay case, the solution only becomes unbounded if time t is allowed to approach $-\infty$.

Example 5.20: Simple Growth Model

Suppose \$5,000 is deposited into an account which earns continuously compounded interest. Under these conditions, the balance in the account grows at a rate proportional to the current balance. Suppose that after 4 years the account is worth \$7,000.

(a) How much is the account worth after 5 years?

(b) How many years does it take for the balance to double?

Solution. Let $y(t)$ denote the balance in the account at the start of year t. Then
$$\frac{dy}{dt} = ky(t),$$
for some constant k. We can solve this differential equation using separation of variables to obtain
$$y(t) = y_0 e^{kt} = 5000e^{kt},$$
where we used the fact that $y(0) = 5,000$. We know that $y(4) = 7000$, which we now use to solve for k:
$$\begin{aligned} 7000 &= 5000e^{4k} \\ e^{4k} &= \frac{7}{5} \\ k &= \frac{\ln\left(\frac{7}{5}\right)}{4} \end{aligned}$$

Therefore,

$$y(t) = 5000e^{\frac{\ln(7/5)}{4}t}.$$

(a) After 5 years, the balance in the account is

$$y(5) = 5000e^{\frac{\ln(7/5)}{4}\cdot 5} \approx \$7614.30.$$

(b) We wish to find t so that $y(t) = 10000$.

$$\begin{aligned} 10000 &= 5000e^{\frac{\ln(7/5)}{4}t} \\ e^{\frac{\ln(7/5)}{4}t} &= 2 \\ t &= \frac{\ln(16)}{\ln\left(\frac{7}{5}\right)} \approx 8.24 \end{aligned}$$

Thus, it takes just over 8 years for the balance in the account to double in value.

♣

5.2.4. Logistic Growth Model

The simple growth model is unrealistic because a quantity that represents something from real life, say, population, does not grow unrestricted. Typically, food resources, competition, predators, or diseases, to name but a few factors, influence the growth of the population, and how much of the population can be sustained in such an environment.

A more realistic model is the so-called **logistic growth model**, which mimics that as a quantity is growing other factors will influence the growth and slow it down until a certain maximum size is being approached. For example, if a population is growing, then food may become scarce or diseases may break out among the population, and the population growth slows down until a certain sustainable size is reached. Replacing k in the simple model with $r(M - y)$ achieves that as the quantity y increases the growth rate decreases, and furthermore, that the maximum population size that is sustained is M. This maximum is referred to as the **carrying capacity**. So the simple model becomes

$$\frac{dy}{dt} = ry(M - y)$$

for some positive r and M.

In other words, the rate of growth of the quantity y is proportional to both itself and the remaining carrying capacity that the quantity can still grow to. As usual, r is the **growth rate constant**.

To solve this first order non-linear differential equation, notice that the equation is separable

$$\frac{dy}{y(M - y)} = r\,dt.$$

Integrating both sides we obtain

$$\int \frac{dy}{y(M-y)} = \int r\,dt$$

$$\int \left(\frac{1}{y} + \frac{1}{M-y}\right) dy = \int rM\,dt$$

$$\ln|y| - \ln|M-y| = rMt + C$$

$$\ln\left|\frac{M-y}{y}\right| = -rMt - C$$

$$\left|\frac{M-y}{y}\right| = e^{-rMt-C}$$

$$\frac{M-y}{y} = Ae^{-rMt},$$

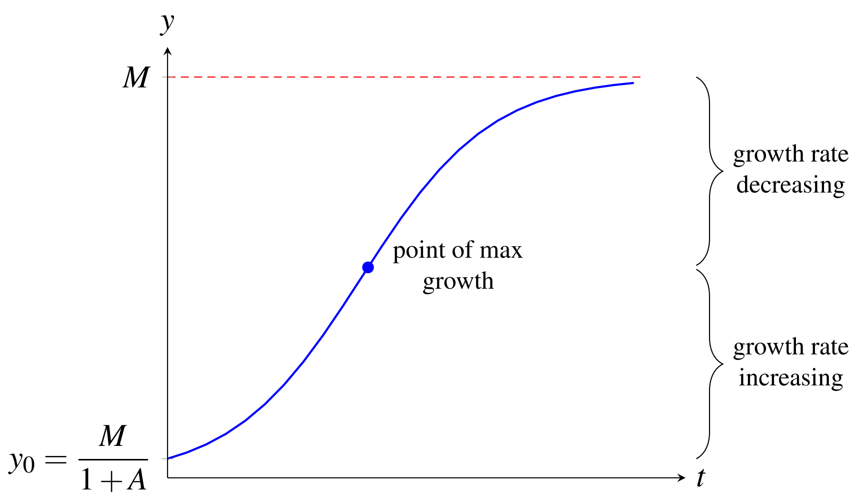
where $A = \dfrac{M-y_0}{y_0}$ for $t_0 = 0$.

Lastly, we solve for y and get the solution

$$y = \frac{M}{1+Ae^{-rMt}} \quad \text{with} \quad A = \frac{M-y_0}{y_0}$$

for some constant y_0 at time $t_0 = 0$.

The graph below shows the typical S-shape of such a solution for some y_0 between 0 and M. For $y_0 > M$, the solution decays exponentially to M. And for $y_0 = M$, the solution remains constant.



Note: Let us rewrite

$$\frac{dy}{dt} = ry(M-y)$$

as

$$\frac{dy}{dt} = rM\left(\frac{M-y}{M}\right)y.$$

Then we can make the following interpretations.

1. When the quantity y is small, then the term $\frac{M-y}{M}$ is close in value to one, and so the differential equation
$$\frac{dy}{dt} = ry(M-y) \approx rMy.$$
In other words, the growth is exponential.

2. However, when the quantity y is near that of the carrying capacity M, then the term $\frac{M-y}{M}$ is close in value to zero. Hence, the less carrying capacity that remains the more the growth rate is slowed down.

Logistic Growth Model

The differential equation
$$\frac{dy}{dt} = ry(M-y)$$
with positive growth constant r and carrying capacity M models **logistic growth** of a quantity y at time t with solution
$$y = \frac{M}{1+Ae^{-rMt}},$$
where $A = \frac{M-y_0}{y_0}$ at time $t_0 = 0$.

Exercises for Section 5.2

Exercise 5.2.1 *Which of the following equations are separable?*

(a) $y' = \sin(ty)$

(b) $y' = e^t e^y$

(c) $yy' = t$

(d) $y' = (t^3 - t)\arcsin(y)$

(e) $y' = t^2 \ln y + 4t^3 \ln y$

Exercise 5.2.2 *Identify the constant solutions (if any) of the following differential equations.*

(a) $y' = t\sin y$

(b) $y' = te^y$

Exercise 5.2.3 *Solve the following differential equations. You may leave your solution in implicit form: that is, you may stop once you have done the integration, without solving for y.*

(a) $y' = 1/(1+t^2)$

(b) $y' = \ln t$

(c) $y' = t/y$

(d) $y' = y^2 - 1$

(e) $y' = t/(y^3 - 5)$

(f) $y' = k(M - y)$

Exercise 5.2.4 *Solve the following initial value problems.*

(a) $y' = t^n$, $y(0) = 1$ *and* $n \geq 0$

(b) $y' = y^{1/3}$, $y(0) = 0$

(c) $y' = ky$, $y(0) = 2$, *and* $y'(0) = 3$

Exercise 5.2.5 *After 10 minutes in Jean-Luc's room, his tea has cooled to* $40°$ *Celsius from* $100°$ *Celsius. The room temperature is* $25°$ *Celsius. How much longer will it take to cool to* $35°$*?*

Exercise 5.2.6 *A radioactive substance obeys the equation* $y' = ky$ *where* $k < 0$ *and* y *is the mass of the substance at time* t*. Suppose that initially, the mass of the substance is* $y(0) = M > 0$*. At what time does half of the mass remain? (This is known as the half life. Note that the half life depends on* k *but not on* M*.)*

Exercise 5.2.7 *Bismuth-210 has a half life of five days. If there is initially 600 milligrams, how much is left after 6 days? When will there be only 2 milligrams left?*

Exercise 5.2.8 *The half life of carbon-14 is 5730 years. If one starts with 100 milligrams of carbon-14, how much is left after 6000 years? How long do we have to wait before there is less than 2 milligrams?*

Exercise 5.2.9 *A certain species of bacteria doubles its population (or its mass) every hour in the lab. The differential equation that models this phenomenon is* $y' = ky$*, where* $k > 0$ *and* y *is the population of bacteria at time* t*. What is* y*?*

Exercise 5.2.10 *If a certain microbe doubles its population every 4 hours and after 5 hours the total population has mass 500 grams, what was the initial mass?*

Exercise 5.2.11 *Given the logistic equation* $y' = ky(M - y)$*,*

(a) *Solve the differential equation for* y *in terms of* t*.*

(b) *Sketch the graph of the solution to this equation when* $M = 1000$*,* $k = 0.002$*,* $y(0) = 1$*.*

Exercise 5.2.12 *The biologist G. F. Gause studied the growth of the protozoan Paramecium in the early 1930s. Through his data, he figured out that the relative growth rate is 0.7944 when* $y(0) = 2$*, and the carrying capacity is 64. This leads to the logistic model*

$$\frac{dy}{dt} = 0.7944\left(1 - \frac{y}{64}\right), \quad y(0) = 2,$$

where time is measured in days.

(a) Solve the differential equation for y in terms of t.

(b) How long will it take for the protozoa to reach 30?

(c) Sketch the graph of the solution to this equation.

5.3 First Order Linear Differential Equations

5.3.1. Homogeneous DEs

A simple, but important and useful, type of separable equation is the **first order homogeneous linear equation**:

Definition 5.21: First Order Homogeneous Linear DE

A **first order homogeneous linear differential equation** *is one of the form $y' + p(t)y = 0$ or equivalently $y' = -p(t)y$.*

We have already seen a first order homogeneous linear differential equation, namely the simple growth and decay model $y' = ky$.

Since first order homogeneous linear equations are separable, we can solve them in the usual way:

$$\begin{aligned} y' &= -p(t)y \\ \int \frac{1}{y}\,dy &= \int -p(t)\,dt \\ \ln|y| &= P(t) + C \\ y &= \pm e^{P(t)+C} \\ y &= Ae^{P(t)}, \end{aligned}$$

where $P(t)$ is an antiderivative of $-p(t)$. As in previous examples, if we allow $A = 0$ we get the constant solution $y = 0$.

Example 5.22: Solving an IVP I

Solve the initial value problem

$$y' + y\cos t = 0,$$

subject to

(a) $y(0) = 1/2$ *(b)* $y(2) = 1/2$

Solution. We start with

$$P(t) = \int -\cos t\,dt = -\sin t,$$

so the general solution to the differential equation is

$$y = Ae^{-\sin t}.$$

(a) To compute the constant coefficient A, we substitute:

$$\frac{1}{2} = Ae^{-\sin 0} = A,$$

so the solutions is

$$y = \frac{1}{2}e^{-\sin t}.$$

(b) To compute the constant coefficient A, we substitute:

$$\begin{aligned} \frac{1}{2} &= Ae^{-\sin 2} \\ A &= \frac{1}{2}e^{\sin 2} \end{aligned}$$

so the solution is

$$y = \frac{1}{2}e^{\sin 2}e^{-\sin t}.$$

♣

Example 5.23: Solving an IVP II

Solve the initial value problem $ty' + 3y = 0$, $y(1) = 2$, *assuming* $t > 0$.

Solution. We write the equation in standard form: $y' + 3y/t = 0$. Then

$$P(t) = \int -\frac{3}{t}\,dt = -3\ln t$$

and

$$y = Ae^{-3\ln t} = At^{-3}.$$

Substituting to find A: $2 = A(1)^{-3} = A$, so the solution is $y = 2t^{-3}$. ♣

5.3.2. Non-Homogeneous DEs

As you might guess, a **first order non-homogeneous linear differential equation** has the form $y' + p(t)y = f(t)$. Not only is this closely related in form to the first order homogeneous linear equation, we can use what we know about solving homogeneous equations to solve the general linear equation.

Definition 5.24: First Order Non-Homogeneous Linear DE

A **first order non-homogeneous linear differential equation** *is one of the form*

$$y' + p(t)y = f(t).$$

Note: When the coefficient of the first derivative is one in the first order non-homogeneous linear differential equation as in the above definition, then we say the DE is in **standard form**.

Let us now discuss how we can find all solutions to a first order non-homogeneous linear differential equation. Suppose that $y_1(t)$ and $y_2(t)$ are solutions to $y' + p(t)y = f(t)$. Let $g(t) = y_1 - y_2$. Then

$$\begin{aligned} g'(t) + p(t)g(t) &= y_1' - y_2' + p(t)(y_1 - y_2) \\ &= (y_1' + p(t)y_1) - (y_2' + p(t)y_2) \\ &= f(t) - f(t) = 0. \end{aligned}$$

In other words, $g(t) = y_1 - y_2$ is a solution to the homogeneous equation $y' + p(t)y = 0$. Turning this around, any solution to the linear equation $y' + p(t)y = f(t)$, call it y_1, can be written as $y_2 + g(t)$, for some particular y_2 and some solution $g(t)$ of the homogeneous equation $y' + p(t)y = 0$. Since we already know how to find all solutions of the homogeneous equation, finding just one solution to the equation $y' + p(t)y = f(t)$ will give us all of them.

Theorem 5.25: General Solution of First Order Non-Homogeneous Linear DE

Given a first order non-homogeneous linear differential equation

$$y' + p(t)y = f(t),$$

let $h(t)$ be a particular solution, and let $g(t)$ be the general solution to the corresponding homogeneous DE

$$y' + p(t)y = 0.$$

Then the **general solution** *to the non-homogeneous DE is constructed as the sum of the above two solutions:*

$$y(t) = g(t) + h(t).$$

Variation of Parameters

We now introduce the first one of two methods discussed in these notes to solve a first order non-homogeneous linear differential equation. Again, it turns out that what we already know helps. We know that the general solution to the homogeneous equation $y' + p(t)y = 0$ looks like $Ae^{P(t)}$, where $P(t)$ is an antiderivative of $-p(t)$. We now make an inspired guess: Consider the function $v(t)e^{P(t)}$, in which we have replaced the constant parameter A with the function $v(t)$. This technique is called **variation of parameters**. For convenience write this as $s(t) = v(t)h(t)$, where $h(t) = e^{P(t)}$ is a solution to the homogeneous equation. Now let's compute a bit with $s(t)$:

$$\begin{aligned} s'(t) + p(t)s(t) &= v(t)h'(t) + v'(t)h(t) + p(t)v(t)h(t) \\ &= v(t)(h'(t) + p(t)h(t)) + v'(t)h(t) \\ &= v'(t)h(t). \end{aligned}$$

The last equality is true because $h'(t) + p(t)h(t) = 0$, since $h(t)$ is a solution to the homogeneous equation. We are hoping to find a function $s(t)$ so that $s'(t) + p(t)s(t) = f(t)$; we will have such a function if we can arrange to have $v'(t)h(t) = f(t)$, that is, $v'(t) = f(t)/h(t)$. But this is as easy (or hard) as finding an antiderivative of $f(t)/h(t)$. Putting this all together, the general solution to $y' + p(t)y = f(t)$ is

$$v(t)h(t) + Ae^{P(t)} = v(t)e^{P(t)} + Ae^{P(t)}.$$

Method of Variation of Parameters

Given a first order non-homogeneous linear differential equation

$$y' + p(t)y = f(t),$$

using **variation of parameters** the general solution is given by

$$y(t) = v(t)e^{P(t)} + Ae^{P(t)},$$

where $v'(t) = e^{-P(t)}f(t)$ and $P(t)$ is an antiderivative of $-p(t)$.

Note: The method of variation of parameters makes more sense after taking linear algebra since the method uses determinants. We therefore restrict ourselves to just one example to illustrate this method.

Example 5.26: Solving an IVP Using Variation of Parameters

Find the solution of the initial value problem $y' + 3y/t = t^2$, $y(1) = 1/2$.

Solution. First we find the general solution; since we are interested in a solution with a given condition at $t = 1$, we may assume $t > 0$. We start by solving the homogeneous equation as usual; call the solution g:

$$g = Ae^{-\int (3/t)\,dt} = Ae^{-3\ln t} = At^{-3}.$$

Then as in the discussion, $h(t) = t^{-3}$ and $v'(t) = t^2/t^{-3} = t^5$, so $v(t) = t^6/6$. We know that every solution to the equation looks like

$$y(t) = v(t)t^{-3} + At^{-3} = \frac{t^6}{6}t^{-3} + At^{-3} = \frac{t^3}{6} + At^{-3}.$$

Finally we substitute $y(1) = \frac{1}{2}$ to find A:

$$\begin{aligned} \frac{1}{2} &= \frac{(1)^3}{6} + A(1)^{-3} = \frac{1}{6} + A \\ A &= \frac{1}{2} - \frac{1}{6} = \frac{1}{3}. \end{aligned}$$

The solution is then

$$y = \frac{t^3}{6} + \frac{1}{3}t^{-3}.$$

♣

Integrating Factor

Another common method for solving such a differential equation is by means of an **integrating factor**. In the differential equation $y' + p(t)y = f(t)$, we note that if we multiply through by a function $I(t)$ to get $I(t)y' + I(t)p(t)y = I(t)f(t)$, the left hand side looks like it could be a derivative computed by the Product Rule:

$$\frac{d}{dt}(I(t)y) = I(t)y' + I'(t)y.$$

Now if we could choose $I(t)$ so that $I'(t) = I(t)p(t)$, this would be exactly the left hand side of the differential equation. But this is just a first order homogeneous linear equation, and we know a solution is $I(t) = e^{Q(t)}$, where $Q(t) = \int p(t)\,dt$. Note that $Q(t) = -P(t)$, where $P(t)$ appears in the variation of parameters method and $P'(t) = -p(t)$. Now the modified differential equation is

$$\begin{aligned} e^{-P(t)}y' + e^{-P(t)}p(t)y &= e^{-P(t)}f(t) \\ \frac{d}{dt}(e^{-P(t)}y) &= e^{-P(t)}f(t). \end{aligned}$$

Integrating both sides gives

$$\begin{aligned} e^{-P(t)}y &= \int e^{-P(t)}f(t)\,dt \\ y &= e^{P(t)}\int e^{-P(t)}f(t)\,dt. \end{aligned}$$

Note: If you look carefully, you will see that this is exactly the same solution we found by variation of parameters, because $e^{-P(t)}f(t) = f(t)/h(t)$. Some people find it easier to remember how to use the integrating factor method, rather than variation of parameters. Since ultimately they require the same calculation, you should use whichever of the two methods appeals to you more.

Definition 5.27: Integrating Factor

Given a first order non-homogeneous linear differential equation

$$y' + p(t)y = f(t),$$

the **integrating factor** *is given by*

$$I(t) = e^{\int p(t)\,dt}.$$

Method of Integrating Factor

Given a first order non-homogeneous linear differential equation

$$y' + p(t)y = f(t),$$

follow these steps to determine the general solution $y(t)$ using an **integrating factor**:

1. Calculate the integrating factor $I(t)$.
2. Multiply the standard form equation by $I(t)$.
3. Simplify the left-hand side to
$$\frac{d}{dt}[I(t)y].$$
4. Integrate both sides of the equation.
5. Solve for $y(t)$.

The solution can be compactly written as

$$y(t) = e^{-\int p(t)\,dt}\left[\int e^{\int p(t)\,dt} f(t)\,dt + C\right].$$

Using this method, the solution of the previous example would look just a bit different.

Example 5.28: Solving an IVP Using Integrating Factor

Find the solution of the initial value problem $y' + 3y/t = t^2$, $y(1) = 1/2$.

Solution. Notice that the differential equation is already in standard form. We begin by computing the integrating factor and obtain

$$I(t) = e^{\int \frac{3}{t}\,dt} = e^{3\ln t} = t^3$$

for $t > 0$.

Next, we multiply both sides of the DE with $I(t)$ and get

$$t^3\left[y' + \frac{3}{y}y\right] = t^3t^2$$
$$y^3y' + 3t^2y = t^5$$

which simplifies to

$$\frac{d}{dt}\left[t^3y\right] = t^5.$$

Now we integrate both sides with respect to t and solve for y:

$$\begin{aligned} \int \frac{d}{dt}\left[t^3 y\right] dt &= \int t^5\, dt \\ t^3 y &= \frac{t^6}{6} + C \\ y &= \frac{t^3}{6} + \frac{C}{t^3}. \end{aligned}$$

Lastly, we use the initial value $y(1) = 1/2$ to find C:

$$y(1) = \frac{1^3}{6} + \frac{C}{1^3} = \frac{1}{2} \implies C = \frac{1}{3}.$$

Hence, the solution to the DE is

$$y = \frac{t^3}{6} + \frac{1}{3t^2}.$$

♣

Example 5.29: General Solution Using Integrating Factor

Determine the general solution of the differential equation

$$\frac{dy}{dt} + 3t^2 y = 6t^2.$$

Solution. We see that the differential equation is in standard form. We then compute the integrating factor as

$$I(t) = e^{\int 3t^2\, dt} = e^{t^3},$$

where we took the arbitrary constant of integration to be zero.

Therefore, we can write the DE as

$$\begin{aligned} e^{t^3}\left[y' + 3t^2 y\right] &= 6t^2 e^{t^3} \\ \frac{d}{dt}\left[e^{t^3} y\right] &= 6t^2 e^{t^3}. \end{aligned}$$

Integrating both sides with respect to t gives

$$e^{t^3} y = 6 \int t^2 e^{t^3}\, dt$$

We solve this integral by making the substitution $u = t^3$, $du = 3t^2\, dt$:

$$\int t^2 e^{t^3}\, dt = \frac{1}{3}\int e^u\, du = \frac{1}{3}e^u + C = \frac{1}{3}e^{t^3} + C.$$

Thus,

$$\begin{aligned} e^{t^3} y &= 2e^{t^3} + C \\ y &= 2 + Ce^{-t^3}. \end{aligned}$$

The general solution to the DE is therefore

$$y = 2 + Ce^{-t^3},$$

for $C \in \mathbb{R}$. ♣

Exercises for Section 5.3

Exercise 5.3.1 *Find the general solution of the following homogeneous differential equations.*

(a) $y' + 5y = 0$

(b) $y' - 2y = 0$

(c) $y' + \dfrac{y}{1+t^2} = 0$

(d) $y' + t^2 y = 0$

Exercise 5.3.2 *Solve the following initial value problems concerning homogeneous DEs.*

(a) $y' + y = 0$, $y(0) = 4$

(b) $y' - 3y = 0$, $y(1) = -2$

(c) $y' + y\sin t = 0$, $y(\pi) = 1$

(d) $y' + ye^t = 0$, $y(0) = e$

(e) $y' + y\sqrt{1+t^4} = 0$, $y(0) = 0$

(f) $y' + y\cos(e^t) = 0$, $y(0) = 0$

(g) $ty' - 2y = 0$, $y(1) = 4$

(h) $t^2 y' + y = 0$, $y(1) = -2$, $t > 0$

(i) $t^3 y' = 2y$, $y(1) = 1$, $t > 0$

(j) $t^3 y' = 2y$, $y(1) = 0$, $t > 0$

Exercise 5.3.3 *Find the general solution of the following non-homogeneous differential equations.*

(a) $y' + 4y = 8$

(b) $y' - 2y = 6$

(c) $y' + ty = 5t$

(d) $y' + e^t y = -2e^t$

(e) $y' - y = t^2$

(f) $2y' + y = t$

Exercise 5.3.4 *Find the general solution of the following non-homogeneous differential equations on the restricted domain.*

(a) $ty' - 2y = 1/t$, $t > 0$

(b) $ty' + y = \sqrt{t}$, $t > 0$

(c) $y'\cos t + y\sin t = 1$, $-\pi/2 < t < \pi/2$

(d) $y' + y\sec t = \tan t$, $-\pi/2 < t < \pi/2$

Exercise 5.3.5 *Solve the following initial value problems concerning non-homogeneous DEs.*

(a) $y' + 4y = 8, y(0) = 1$

(b) $y' - 2y = 6, y(0) = 3$

(c) $y' + ty = 5t, y(2) = 1$

(d) $y' + e^t y = -2e^t, y(0) = e^{-1}$

(e) $y' - y = t^2, y(0) = 4$

(f) $2y' + y = t, y(1) = -1$

Exercise 5.3.6 *A function $y(t)$ is a solution of $y' + ky = 0$. Suppose that $y(0) = 100$ and $y(2) = 4$. Find k and find $y(t)$.*

Exercise 5.3.7 *A function $y(t)$ is a solution of $y' + t^k y = 0$. Suppose that $y(0) = 1$ and $y(1) = e^{-13}$. Find k and find $y(t)$.*

Exercise 5.3.8 *A bacterial culture grows at a rate proportional to its population. If the population is one million at $t = 0$ and 1.5 million at $t = 1$ hour, find the population as a function of time.*

Exercise 5.3.9 *A radioactive element decays with a half-life of 6 years. If a mass of the element weighs ten pounds at $t = 0$, find the amount of the element at time t.*

5.4 Approximation

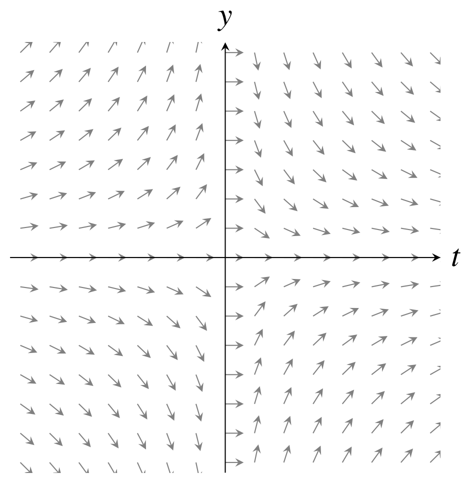
We have seen how to solve a restricted collection of differential equations, or more accurately, how to attempt to solve them—we still may not be able to find the required antiderivatives. Not surprisingly, non-linear equations can be even more difficult to solve. Yet much is known about solutions to some more general equations.

Suppose $\phi(t, y)$ is a function of two variables. A more general class of first order differential equations has the form $y' = \phi(t, y)$. This is not necessarily a linear first order equation, since ϕ may depend on y in some complicated way; note however that y' appears in a very simple form. Under suitable conditions on the function ϕ, it can be shown that every such differential equation has a solution, and moreover that for each initial condition the associated initial value problem has exactly one solution. In practical applications this is obviously a very desirable property.

5.4.1. Slope Field

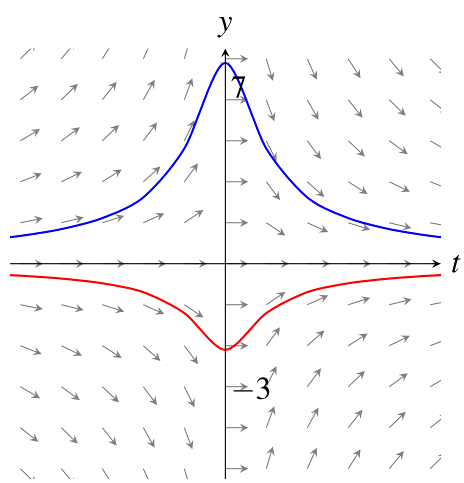
There are many ways to visualize solutions of differential equations of the form $y' = \phi(t, y)$. A particularly simple and useful way to do so is using a so-called **slope field**, where we make use of the fact that $\phi(t, y)$ is just the slope of the solution at any point (t, y). This allows us to generate a plot in the t-y-plane of the slope at any point as shown below for the DE

$$\frac{dy}{dt} = -\frac{ty}{1 + t^2}.$$



Notice that any solution curve regardless of its initial value must be tangent to the slope field at any point along the curve. So the slope field gives us a simple way of graphing a curve that approximates a specific solution. Below, we see the solution curves for $y(0) = -3$ and $y(0) = 7$ respectively of the DE

$$\frac{dy}{dt} = -\frac{ty}{1+t^2}.$$



Note: Aside from being able to visually observe the solution to a differential equation, the slope field also provides us with information about the **long-term behaviour** of such solutions as t increases. In the above slope field, both solutions tend to zero as t goes to infinity.

Nowadays, slope fields are generated using a computer algebra system like Sage or Maple rather than by hand. However, for the sake of understanding how to produce slope fields, let us now look at two examples of how to generate such a slope field by hand. We begin with an example, where $\phi(t,y)$ is not dependent on t.

Example 5.30: Generating a Slope Field that is not Time-Dependent

Plot the slope field of the differential equation

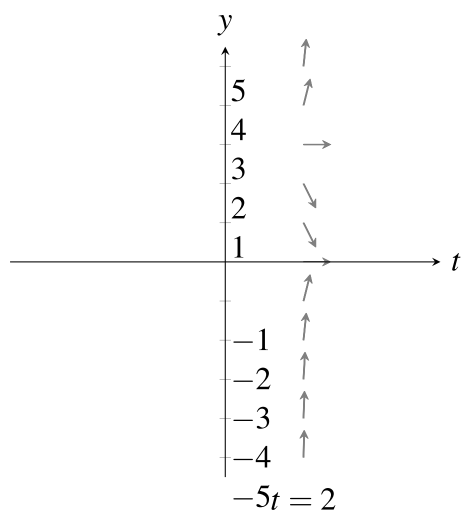
$$\frac{dy}{dt} = y^2 - 3y.$$

Solution. Notice that the right side of the DE is not dependent on t. This means that if we plot the slopes along a vertical line of the t-y-plane, we can produce the full slope field by simply copying these slopes to the left and right as t decreases and increases.

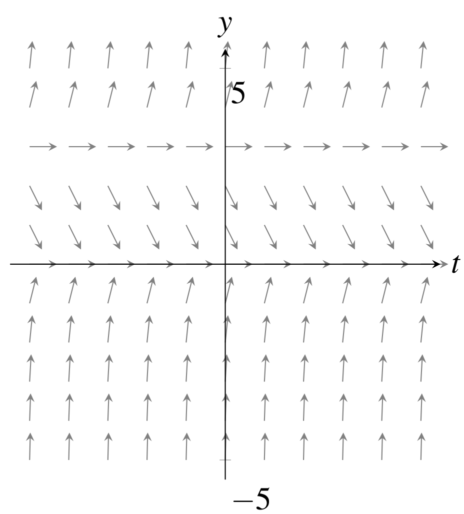
We begin by computing the slopes for y ranging between -5 and 5 in step-sizes of one:

y	-5	-4	-3	-2	-1	0	1	2	3	4	5
y'	40	28	18	10	4	0	-2	-2	0	4	10

Next, we graph short tangent line segments at the points (t,y) for, say, $t = 2$:



Lastly, we produce the full slope field by making copies of this vertical slice of slopes to the left and right as shown below for values of t between -5 and 5:



♣

Let us now look at an example, where $\phi(t,y)$ depends on t and y, in which case the computations are more labourious.

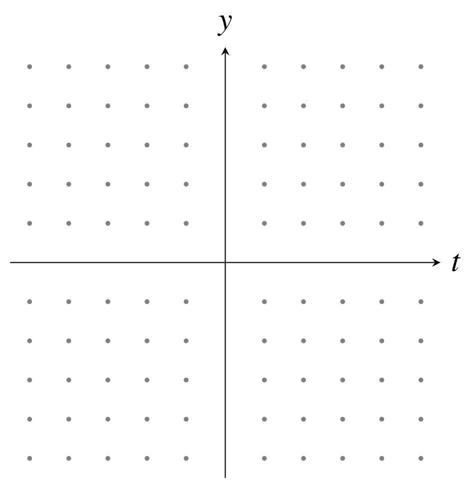
Example 5.31: Generating a Slope Field that is Time-Dependent

Plot the slope field of the differential equation

$$\frac{dy}{dt} = t - y^2.$$

Solution. Notice that the right side of the differential equation is dependent on both t and y. We can therefore not replicate slopes as showcased in the above example, but must employ a different method. In fact, we demonstrate two methods, one that is crude but will always work, and one where we identify places where the derivative will be constant.

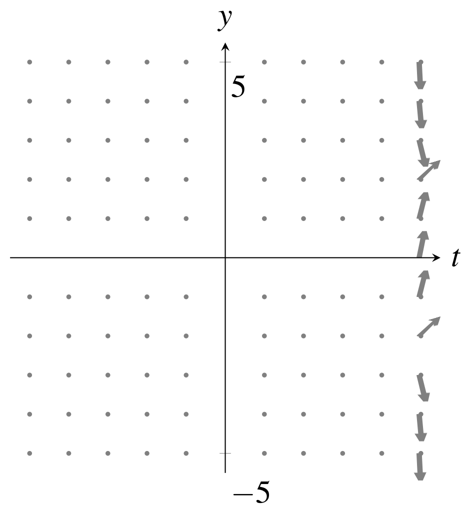
Method 1: In this approach, we simply compute the slope values at integer coordinates for a certain section of the t-y-plane, say the $[-5,5] \times [-5,5]$ grid shown below.



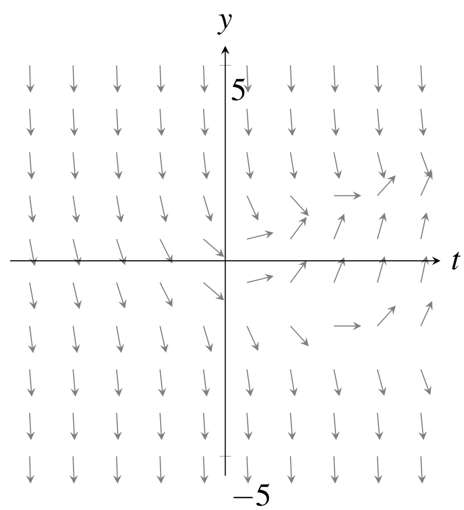
Before we start computing slope values, we investigate the right side of the differential equation and take note of any symmetry that might reduce the amount of computations that we need to do. Here, we notice the symmetry in y, that is, the slope for a certain y-value is the same as for the corresponding negative y-value at the same value of t. This reduces our computations by half. Let's begin calculating the slope values for $-5 \le y \le 5$ when $t = 5$:

y	-5	-4	-3	-2	-1	0	1	2	3	4	5
y'	-20	-11	-4	1	4	5	4	1	-4	-11	-20

Next, we draw short tangent line segments at the grid points $(5,y)$ as before to get the graph as shown:



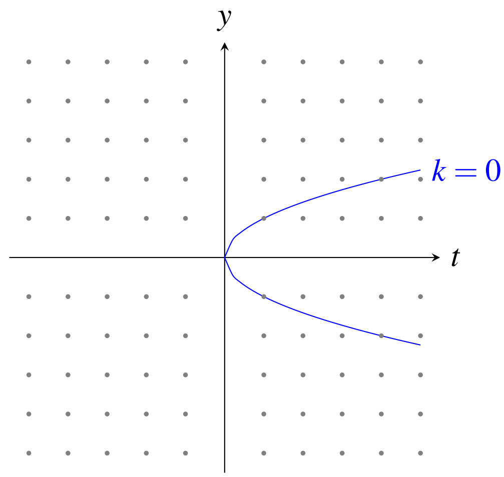
We repeat the computations for all integer values of t and y on the $[-5,5] \times [-5,5]$ grid to obtain the slope field below.



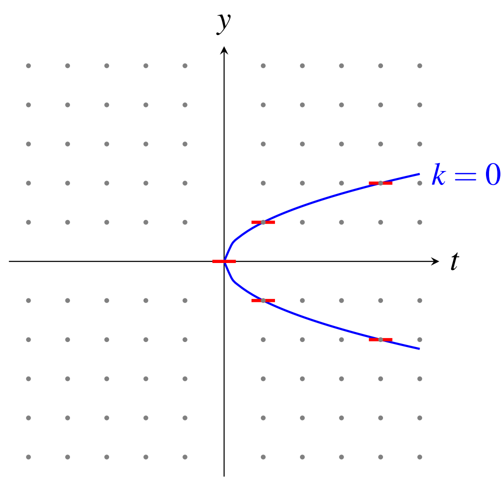
Method 2: In this approach, we look for places where the derivative will be constant, that is, where

$$t - y^2 = k,$$

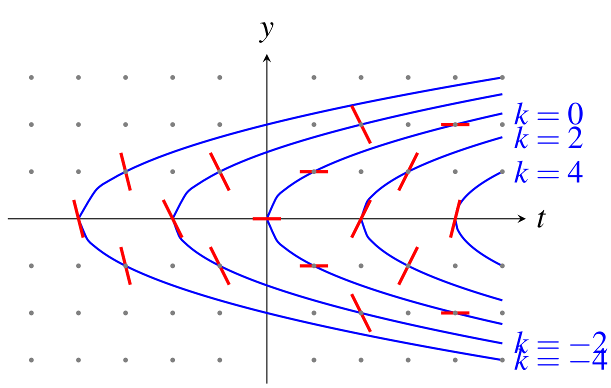
for some constant k. When $k = 0$, then this produces the curve $y^2 = t$, which is the parabola shown below:



At all points along this curve, the slope of the differential equation must be zero (see below).



In fact, all curves $y^2 = t - k$ are parabolas shifted to the left or right depending on the value of k, and the slopes along any of these curves must have the value k, which is shown in the following graph.



Ignoring the parabolas on the above plot, we have generated the desired slope field. ♣

Note:

1. The curves produced by the equation $\phi(t,y) = k$ for k constant and any DE $y' = \phi(t,y)$, as in the above example with $t - y^2 = k$, are referred to as a **family of curves**. They arise with differential equations but also other areas of mathematics.

2. Method 1 in Example 5.31 does not depend on an integer grid. One can choose any points to compute the slopes.

We now provide one example that outlines how to draw the solution curve to a differential equation on a slope field for a specific initial value. We point out that in this process, we often make use of the simplest slope, namely the zero slope, as it eases computations.

Example 5.32: Graphing Solution Curves on a Slope Field

Given the slope field from Example 5.31 for the differential equation

$$\frac{dy}{dt} = t - y^2,$$

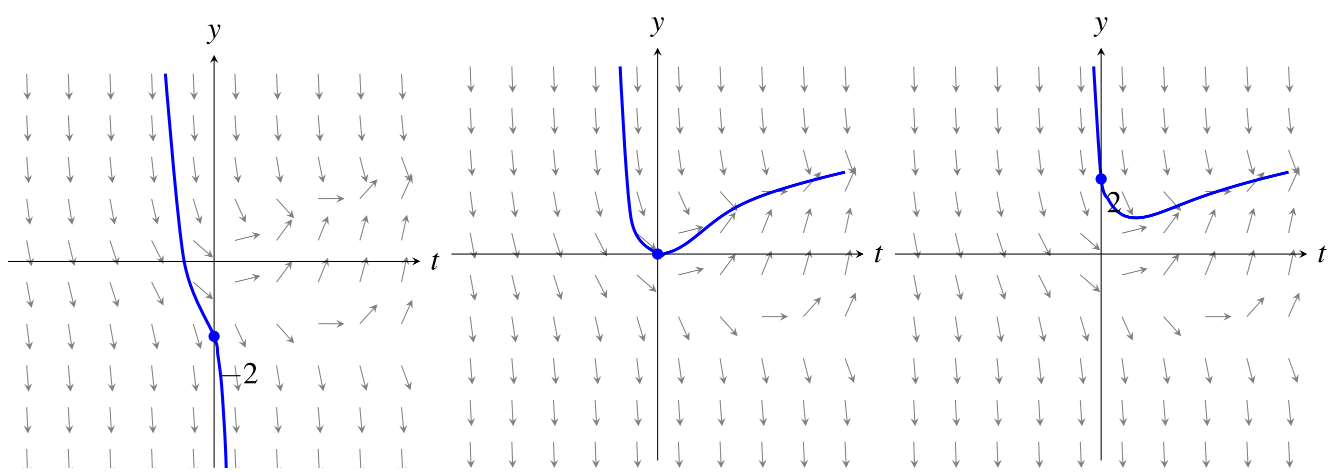
graph the solution curves for the following initial values:

(a) $y(0) = -2$ *(b)* $y(0) = 0$ *(c)* $y(0) = 2$

Solution.

(a) We start by indicating the point $y(0) = -2$ and then follow the solution trajectory forward and backward in the independent variable t based on the little tangent lines. The solution curve is shown below to the left.

(b) and (c) A similar approach as in part (a) produces the solution curves when $y(0) = 0$ shown below centre and when $y(0) = 2$ shown below to the right.



(a) (b) (c)

♣

Even when a differential equation can be solved explicitly, the slope field can help in understanding what the solutions look like with various initial conditions. Recall the logistic equation $y' = ky(M-y)$: y is a population at time t, M is a measure of how large a population the environment can support, and k measures the reproduction rate of the population. Figure 5.1 shows a slope field for this equation that is quite informative. It is apparent that if the initial population is smaller than M it rises to M over the long term, while if the initial population is greater than M it decreases to M.

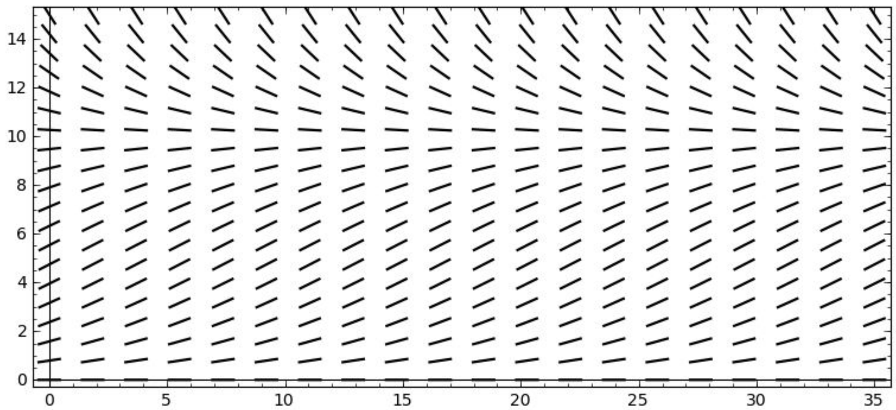
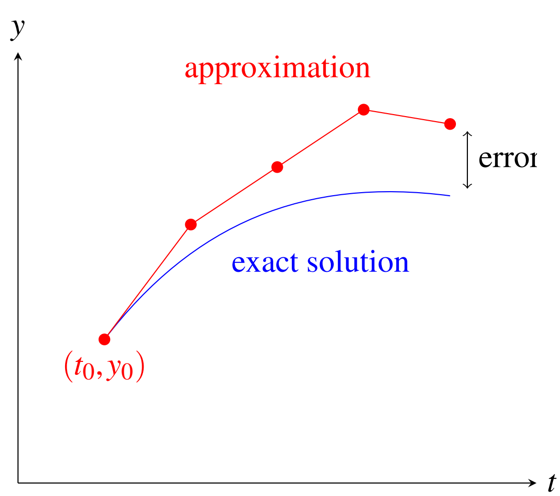


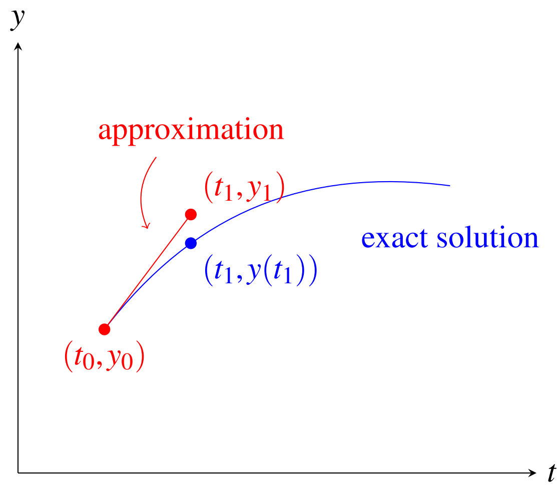
Figure 5.1: A slope field for $y' = 0.2y(10-y)$**.**

5.4.2. Euler's Method

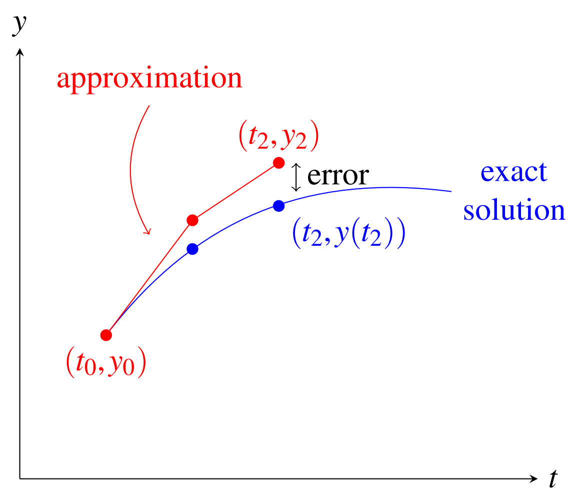
Not all differential equations that are important in practice can be solved exactly, so techniques have been developed to numerically approximate solutions. Such techniques are referred to as **numerical methods**, which are often quite accurate, and typically executed by means of a computer and so also fast. When exact computation are not possible or excessively complicated, then numerical methods are chosen. We describe one such method, **Euler's Method**, which is simple though not particularly useful compared to some more sophisticated techniques. Euler's method is based on the method of linearization by approximating the solution curve with a succession of linearizations over an interval as shown in the figure below.



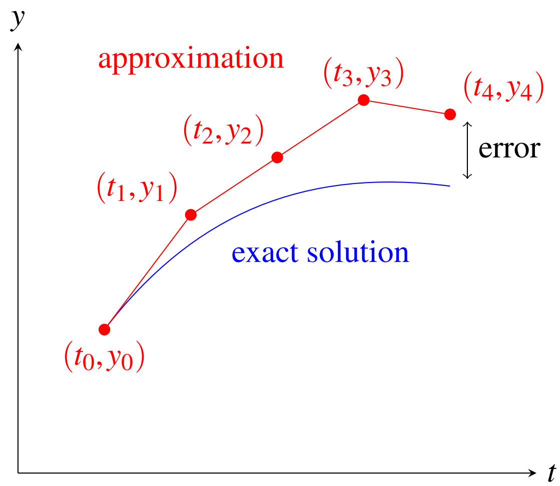
Suppose we wish to approximate a solution to the initial value problem $y' = \phi(t,y)$, $y(t_0) = y_0$, for $t \geq t_0$. Under reasonable conditions on ϕ, we know the solution exists, represented by a curve in the t-y-plane; call this solution $f(t)$ as shown below. The point (t_0, y_0) is of course on this curve. We also know the slope of the curve at this point, namely $\phi(t_0, y_0)$. If we follow the tangent line for a brief distance, we arrive at a point that should be almost on the graph of $f(t)$, namely $(t_0 + \Delta t, y_0 + \phi(t_0, y_0)\Delta t)$; call this point (t_1, y_1). The graph below depicts the first linearization starting from the initial value (t_0, y_0).



Now we pretend, in effect, that this point (t_1, y_1) really is on the graph of $f(t)$, in which case we again know the slope of the curve through (t_1, y_1), namely $\phi(t_1, y_1)$. So we can compute a new point, $(t_2, y_2) = (t_1 + \Delta t, y_1 + \phi(t_1, y_1)\Delta t)$ that is a little farther along, still close to the graph of $f(t)$ but probably not quite so close as (t_1, y_1). The graph below shows two steps in the Euler approximation and the accumulated error.



We can continue in this way as shown below, doing a sequence of straightforward calculations, until we have an approximation (t_n, y_n) for whatever time t_n we need.



At each step we do essentially the same calculation, namely:

$$(t_{i+1}, y_{i+1}) = (t_i + \Delta t, y_i + \phi(t_i, y_i)\Delta t).$$

Note: We expect that smaller time steps Δt will give better approximations, but of course it will require more work to compute to a specified time. It is possible to compute a guaranteed upper bound on how far off the approximation might be, that is, how far y_n is from $f(t_n)$. Suffice it to say that the bound is not particularly good and that there are other more complicated approximation techniques that do better.

Guideline for Euler's Method

Given a first order differential equation

$$\frac{dy}{dt} = \phi(t, y),$$

with initial condition $y(t_0) = y_0$, the value of $y(t_n)$ for some $t_n > t_0$ and integer n is approximated with Euler's method as follows:
Define the step size

$$h = \frac{t_n - t_0}{n}.$$

For $1 \leq i \leq n$ compute the following sequence of numbers recursively

$$y_i = y_{i-1} + h\phi(t_{i-1}, y_{i-1}).$$

Then the hope is that $y(t_i) \approx y_i$.

Note:

1. The above method is based on a **uniform** step size; however, Euler's method can also be performed using a **non-uniform** step size.

2. Typically, Euler's method is performed using a computer program, as the method is algorithmic.

We first illustrate Euler's method with one example, and then we demonstrate how to use a table if computations are performed by hand rather than a computer.

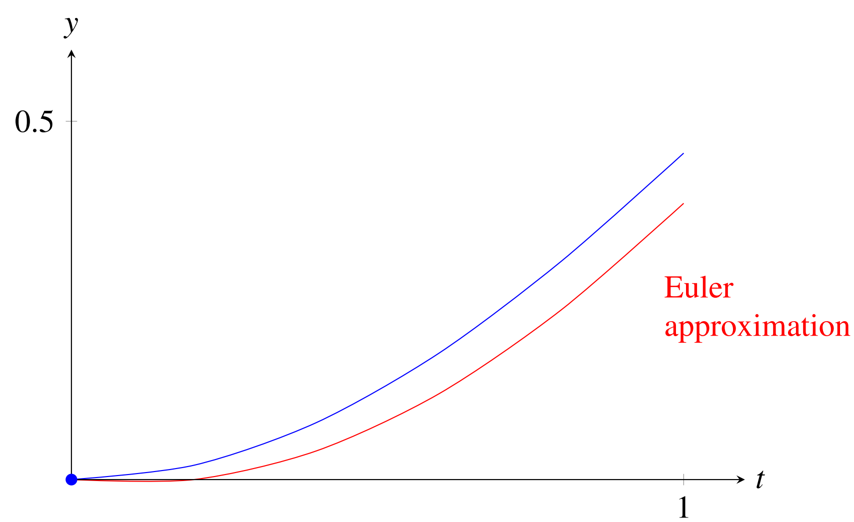
Example 5.33: Approximating a Solution Using Euler's Method

Compute an approximation to the solution for $y' = t - y^2$, $y(0) = 0$, when $t = 1$.

Solution. We will use $\Delta t = 0.2$, which is easy to do even by hand, though we should not expect the resulting approximation to be very good. We get

$$\begin{aligned}
(t_1, y_1) &= (0 + 0.2, 0 + (0 - 0^2)0.2) = (0.2, 0) \\
(t_2, y_2) &= (0.2 + 0.2, 0 + (0.2 - 0^2)0.2) = (0.4, 0.04) \\
(t_3, y_3) &= (0.6, 0.04 + (0.4 - 0.04^2)0.2) = (0.6, 0.11968) \\
(t_4, y_4) &= (0.8, 0.11968 + (0.6 - 0.11968^2)0.2) = (0.8, 0.23681533952) \\
(t_5, y_5) &= (1.0, 0.23681533952 + (0.6 - 0.23681533952^2)0.2) = (1.0, 0.385599038513605)
\end{aligned}$$

So $y(1) \approx 0.3856$. As it turns out, this is not accurate to even one decimal place. The graph below shows these points connected by line segments (the lower curve) compared to a solution obtained by a much better approximation technique. Note that the shape is approximately correct even though the end points are quite far apart.



♣

Note: If you need to do Euler's method by hand, it is useful to construct a table to keep track of the work, as shown in Table 5.1. Each row holds the computation for a single step: The starting point (t_i, y_i); the stepsize Δt; the computed slope $\phi(t_i, y_i)$; the change in y, $\Delta y = \phi(t_i, y_i)\Delta t$; and the new point, $(t_{i+1}, y_{i+1}) = (t_i + \Delta t, y_i + \Delta y)$. The starting point in each row is the newly computed point from the end of the previous row.

(t,y)	Δt	$\phi(t,y)$	$\Delta y = \phi(t,y)\Delta t$	$(t+\Delta t, y+\Delta y)$
$(0,0)$	0.2	0	0	$(0.2,0)$
$(0.2,0)$	0.2	0.2	0.04	$(0.4,0.04)$
$(0.4,0.04)$	0.2	0.3984	0.07968	$(0.6,0.11968)$
$(0.6,0.11968)$	0.2	$0.58\ldots$	$0.117\ldots$	$(0.8,0.236\ldots)$
$(0.8,0.236\ldots)$	0.2	$0.743\ldots$	$0.148\ldots$	$(1.0,0.385\ldots)$

Table 5.1: Computing with Euler's Method.

Note: If we obtain a slope field for $\phi(t,y)$, then one can sketch reasonably accurate solution curves, in essence doing Euler's method visually. A slope field for $y' = t - y^2$ is shown in Figure 5.2.

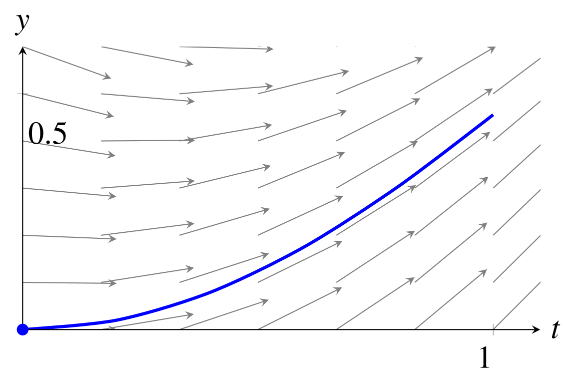
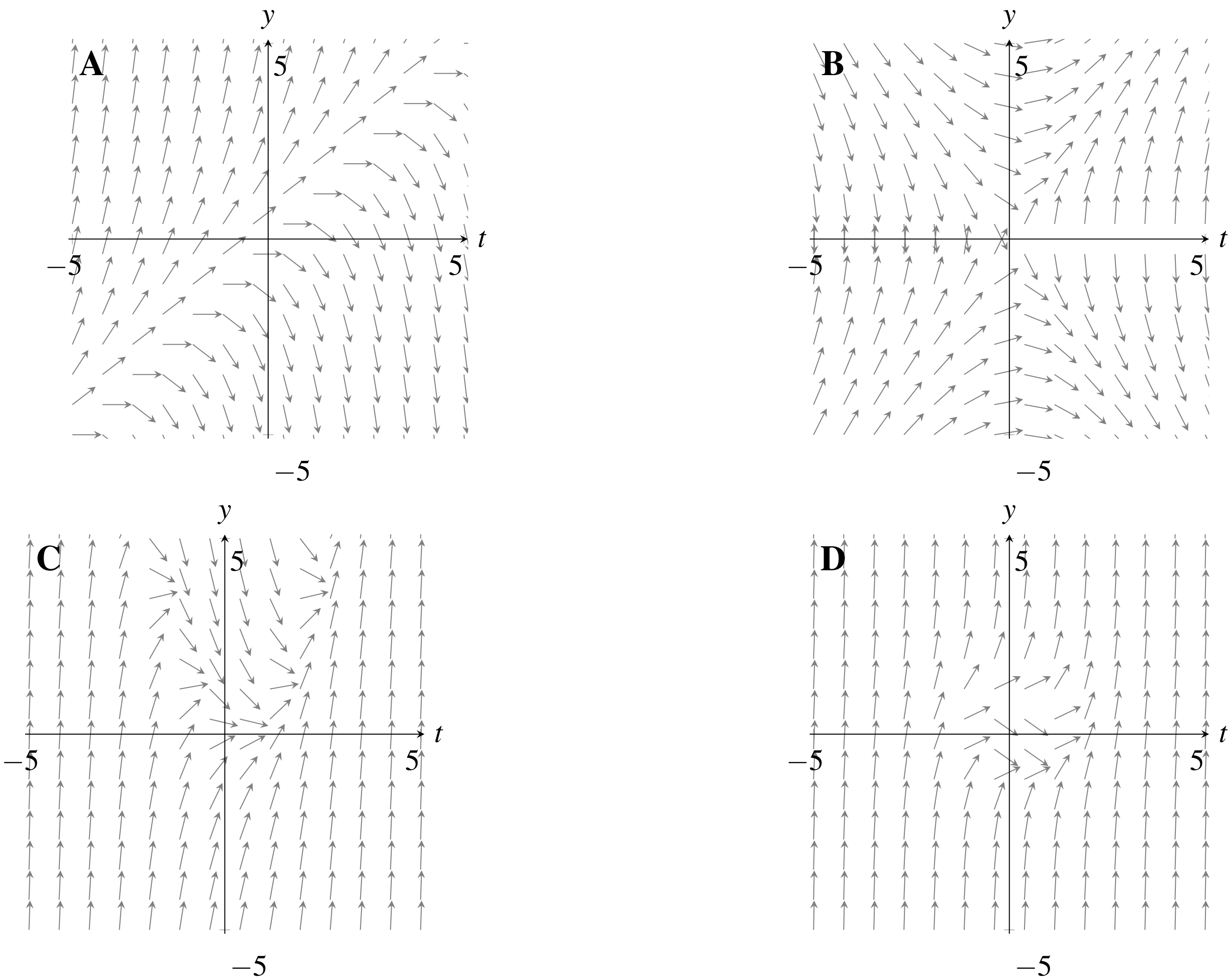


Figure 5.2: A slope field for $y' = t - y^2$, and the solution curve for $y(0) = 0$.

Exercises for Section 5.4

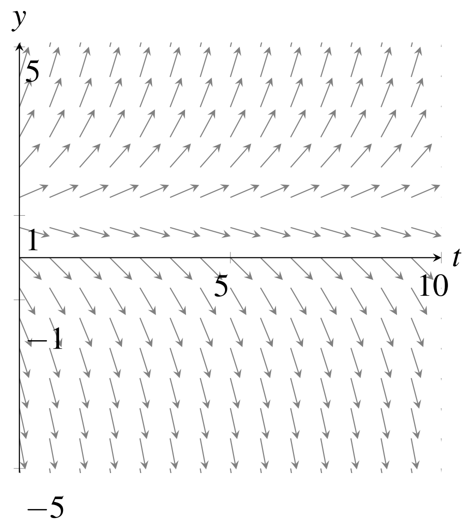
Exercise 5.4.1 *Match the slope fields to the differential equation.*



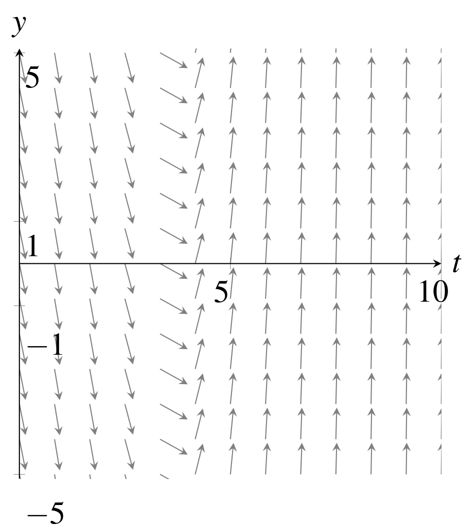
(a) $\dfrac{dy}{dt} = -\dfrac{2t}{y}$ *(b)* $\dfrac{dy}{dt} = y - t$ *(c)* $\dfrac{dy}{dy} = t^2 - y$ *(d)* $\dfrac{dy}{dt} = t^2 + y^2 - 1$

Exercise 5.4.2 *Given the slope field of the differential equation, sketch in the solution curves for* $y(0) = -1$ *and* $y(0) = 1$ *respectively.*

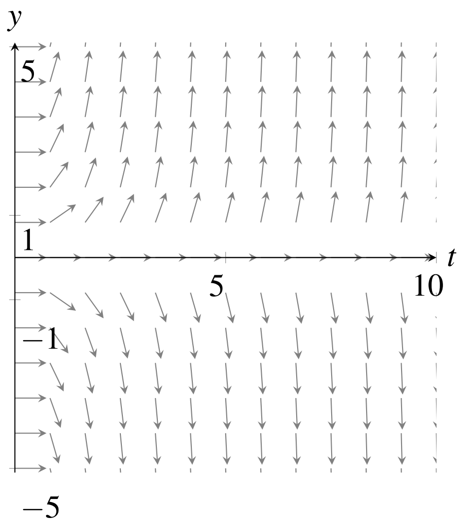
(a) $\dfrac{dy}{dt} = y - 1$



(b) $\dfrac{dy}{dt} = t^2 - 2t - 5$



(c) $\dfrac{dy}{dt} = ty$



Exercise 5.4.3 *Develop the slope field to the differential equation, and add the solution curve with the given initial value.*

(a) $\dfrac{dy}{dt} = y^2,\ y(0) = -2$

(b) $\dfrac{dy}{dt} = \cos(t),\ y(0) = 0$

(c) $\dfrac{dy}{dt} = e^{-y},\ y(0) = 2$

Exercise 5.4.4 *Compute the Euler approximations for the initial value problem for $0 \le t \le 1$ and $\Delta t = 0.2$. If you have access to a computer algebra system like Sage or Maple, generate the slope field first and attempt to sketch the solution curve. Then use the computer algebra system to compute better approximations with smaller values of Δt.*

(a) $y' = t/y$, $y(0) = 1$

(b) $y' = t + y^3$, $y(0) = 1$

(c) $y' = \cos(t+y)$, $y(0) = 1$

(d) $y' = t \ln y$, $y(0) = 2$

6. Sequences and Series

In this Chapter, we introduce the concepts of infinite sequences and series. One important application is the representation of a differentiable function as an infinite sum of powers of the independent variable. This allows us to extend differentiation and integration to a more general class of functions that cannot necessarily be represented in terms of elementary functions like power, trigonomentric, exponential and logarithmic functions.

Consider the following sum:

$$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\cdots+\frac{1}{2^i}+\cdots$$

The dots at the end indicate that the sum goes on forever. Does this make sense? Can we assign a numerical value to an infinite sum? While at first it may seem difficult or impossible, we have certainly done something similar when we talked about one quantity getting "closer and closer" to a fixed quantity. Here we could ask whether, as we add more and more terms, the sum gets closer and closer to some fixed value. That is, look at

$$\begin{aligned}
\frac{1}{2}&=\frac{1}{2}\\
\frac{3}{4}&=\frac{1}{2}+\frac{1}{4}\\
\frac{7}{8}&=\frac{1}{2}+\frac{1}{4}+\frac{1}{8}\\
\frac{15}{16}&=\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}
\end{aligned}$$

and so on, and consider whether these values have a limit. It seems likely that they do, namely 1. In fact, as we will see, it's not hard to show that

$$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\cdots+\frac{1}{2^i}=\frac{2^i-1}{2^i}=1-\frac{1}{2^i}$$

and then

$$\lim_{i\to\infty}\left(1-\frac{1}{2^i}\right)=1-0=1.$$

There is a context in which we already implicitly accept this notion of infinite sum without really thinking of it as a sum: The representation of a real number as an infinite decimal. For example,

$$0.3333\bar{3}=\frac{3}{10}+\frac{3}{100}+\frac{3}{1000}+\frac{3}{10000}+\cdots=\frac{1}{3},$$

or likewise

$$3.14159\ldots=3+\frac{1}{10}+\frac{4}{100}+\frac{1}{1000}+\frac{5}{10000}+\frac{9}{100000}+\cdots=\pi.$$

An infinite sum is called a **series**, and is usually written using the same sigma notation that we encountered in Chapter 1. In this case, however, we use infinity as an upper limit of summation to indicate that there is no 'last term'. The series we first examined can be written as

$$\frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2^i}+\cdots=\sum_{i=1}^{\infty}\frac{1}{2^i}$$

A related notion that will aid our investigations is that of a **sequence**. A sequence is just an ordered (possibly infinite) list of numbers. For example, the terms in the infinite sequence above are an example of a sequence:

$$1,\frac{1}{2},\frac{1}{3},\ldots,\frac{1}{n},\ldots$$

We will begin by learning some useful facts about sequences.

6.1 Sequences

While the idea of a sequence of numbers, $a_1,a_2,a_3,\ldots$ is straightforward, it is useful to think of a sequence as a function. We have dealt with functions whose domains are the real numbers, or a subset of the real numbers, like $f(x)=\sin x$. A sequence can be regarded as a function with domain as the natural numbers $\mathbb{N}=\{1,2,3,\ldots\}$ or the non-negative integers, $\mathbb{Z}^{\geq 0}=\{0,1,2,3,\ldots\}$. The range of the function is still allowed to be the set of all real numbers. We say that a sequence is a function $f:\mathbb{N}\to\mathbb{R}$.

Definition 6.1: Infinite Sequence

*An **infinite sequence** of numbers is a function $f:\mathbb{N}\to\mathbb{R}$ such that*

$$f(n)=a_n$$

where
$a_1,\ a_2,\ a_3,\ldots$ *are the **terms** of the sequence, and*
a_n *is the **n-th term** of the sequence.*

Note: Sequences are commonly denoted in several different, but equally acceptable ways:

$$a_1,a_2,a_3,\ldots$$
$$\{a_n\}_{n=1}^{\infty}$$
$$\{f(n)\}_{n=1}^{\infty}$$

As with functions of the real numbers, we will most often encounter sequences that can be expressed by a formula. We have already seen the sequence $a_i=f(i)=1-1/2^i$. Some other simple examples are:

$$f(i)=\frac{i}{i+1}$$
$$f(n)=\frac{1}{2^n}$$

$$f(n) = \sin(n\pi/6)$$
$$f(i) = \frac{(i-1)(i+2)}{2^i}$$

Frequently these formulas will make sense if thought of either as functions with domain $\mathbb{R}$ or $\mathbb{N}$, though occasionally one will make sense for integer values only.

The main question of interest when dealing with sequences is what happens to the terms as we go further and further down the list. In particular, as i becomes extremely large, does a_i get closer to one specific value? This is reminiscent of a question we asked in Chapter 3 of *Calculus Early Transcendentals Differential & Multi-Variable Calculus for Social Sciences*, when looking at limits of functions. In fact, the problems are closely related and we define the limit of a sequence in a way similar to the definition for *Limit at Infinity* in that chapter.

Definition 6.2: Limit of a Sequence

Suppose that $\{a_n\}_{n=1}^{\infty}$ is an infinite sequence. We write

$$\lim_{n\to\infty} a_n = L,$$

if a_n can be made arbitrarily close to L by taking n large enough. If this limit exists, we say that the sequence **converges** *to L, otherwise it* **diverges**.

Note: Intuitively, $\lim\limits_{n\to\infty} a_n = L$ means that the further we go in the sequence, the closer the terms get to L.

Example 6.3: Exponential Sequence

Show that $\{2^{1/n}\}_{n=1}^{\infty}$ converges to 1.

Solution. We see that $a_n = 2^{1/n}$ and compute the following limit:

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} 2^{1/n} = 2^0 = 1.$$

Therefore, the sequence converges to 1. ♣

If a sequence is defined by a formula $\{f(n)\}_{n=1}^{\infty}$, we can often expand the domain of the function f to the set of all (or almost all) real numbers. For example, $f(n) = \frac{1}{n}$ is defined for all non-zero real numbers.

When this happens, we can sometimes find the limit of the sequence $\{f(n)\}_{n=1}^{\infty}$ more easily by finding the limit of the function $f(x)$, $x \in \mathbb{R}$, as x approaches infinity.

Theorem 6.4: Limit of a Sequence

If $\lim\limits_{x\to\infty} f(x) = L$, where $f : \mathbb{R} \to \mathbb{R}$, then $\{f(n)\}_{n=1}^{\infty}$ converges to L.

Proof. This follows immediately from the definition for Limit at Infinity in Chapter 3 of *Calculus Early Transcendentals Differential & Multi-Variable Calculus for Social Sciences*. ♣

Note: Hereafter we will use the convention that x refers to a real-valued variable and i and n are integer-valued.

Example 6.5: Sequence of $1/n$

Show that $\left\{\frac{1}{n}\right\}_{n=1}^{\infty}$ *converges to 0.*

Solution. Since $\lim\limits_{x\to\infty}\frac{1}{x}=0$, $\lim\limits_{n\to\infty}\frac{1}{n}=0$. ♣

Note:

1. The converse of Theorem 6.4 is not true. Let $f(n)=\sin(n\pi)$. This is the sequence
$$\sin(0\pi),\sin(1\pi),\sin(2\pi),\sin(3\pi),\ldots=0,0,0,0,\ldots$$
since $\sin(n\pi)=0$ when n is an integer. Thus $\lim\limits_{n\to\infty}f(n)=0$. But $\lim\limits_{x\to\infty}f(x)$, when x is real, does not exist: as x gets bigger and bigger, the values $\sin(x\pi)$ do not get closer and closer to a single value, but take on all values between -1 and 1 over and over.

2. In general, whenever you want to know $\lim\limits_{n\to\infty}f(n)$ you should first attempt to compute $\lim\limits_{x\to\infty}f(x)$, since if the latter exists it is also equal to the first limit. But if for some reason $\lim\limits_{x\to\infty}f(x)$ does not exist, it may still be true that $\lim\limits_{n\to\infty}f(n)$ exists, but you'll have to figure out another way to compute it.

It is occasionally useful to think of the graph of a sequence. Since the function is defined only for integer values, the graph is just a sequence of points. In Figure 6.1 we see the graphs of two sequences and the graphs of the corresponding real functions.

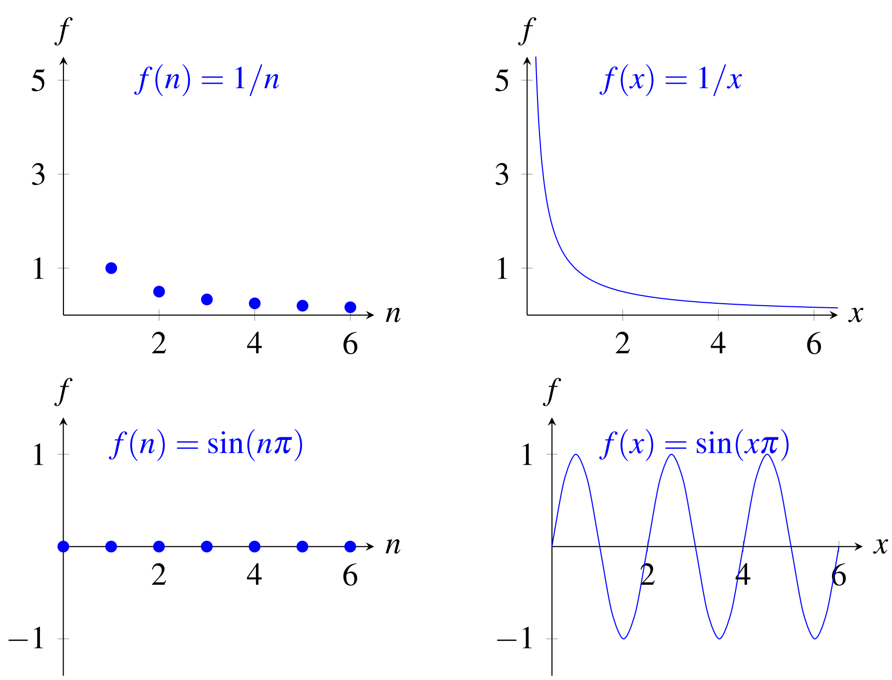


Figure 6.1: Graphs of sequences and their corresponding real functions.

Not surprisingly, the properties of limits of real functions translate into properties of sequences quite easily. The theorem on Limit Properties in Chapter 3 of *Calculus Early Transcendentals Differential & Multi-Variable Calculus for Social Sciences* becomes:

Theorem 6.6: Properties of Sequences

Suppose that $\lim_{n\to\infty} a_n = L$ *and* $\lim_{n\to\infty} b_n = M$ *and* k *is some constant. Then*

$$\lim_{n\to\infty} ka_n = k\lim_{n\to\infty} a_n = kL$$

$$\lim_{n\to\infty} (a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n = L + M$$

$$\lim_{n\to\infty} (a_n - b_n) = \lim_{n\to\infty} a_n - \lim_{n\to\infty} b_n = L - M$$

$$\lim_{n\to\infty} (a_n b_n) = \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n = LM$$

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{\lim_{n\to\infty} a_n}{\lim_{n\to\infty} b_n} = \frac{L}{M}, \text{ if } M \text{ is not } 0$$

Likewise the Squeeze Theorem in Chapter 3 of *Calculus Early Transcendentals Differential & Multi-Variable Calculus for Social Sciences* becomes:

Theorem 6.7: Squeeze Theorem for Sequences

Suppose that $a_n \le b_n \le c_n$ *for all* $n > N$, *for some* N. *If* $\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = L$, *then* $\lim_{n\to\infty} b_n = L$.

And a final useful fact:

Theorem 6.8: Absolute Value Sequence

$\lim_{n\to\infty} |a_n| = 0$ *if and only if* $\lim_{n\to\infty} a_n = 0$.

Note: This says simply that the size of $|a_n|$ gets close to zero if and only if a_n gets close to zero.

Example 6.9: Convergence of a Rational Function

Determine whether $\left\{ \frac{n}{n+1} \right\}_{n=0}^{\infty}$ *converges or diverges. If it converges, compute the limit.*

Solution. Defining $f(x) = \frac{x}{x+1}$ we obtain

$$\lim_{x\to\infty} \frac{x}{x+1} = \lim_{x\to\infty} 1 - \frac{1}{x+1} = 1 - 0 = 1.$$

Thus the sequence converges to 1. ♣

Example 6.10: Convergence of Ratio with Natural Logarithm

Determine whether $\left\{\frac{\ln n}{n}\right\}_{n=1}^{\infty}$ *converges or diverges. If it converges, compute the limit.*

Solution. We define $f(x)=\frac{\ln x}{x}$ and compute

$$\lim_{x\to\infty}\frac{\ln x}{x}=\lim_{x\to\infty}\frac{1/x}{1}=0,$$

using L'Hôpital's Rule. Thus the sequence converges to 0. ♣

Example 6.11: Alternating Terms

Determine whether $\{(-1)^n\}_{n=0}^{\infty}$ *converges or diverges. If it converges, compute the limit.*

Solution. $f(x)=(-1)^x$ is undefined for irrational values of x so $\lim\limits_{x\to\infty}(-1)^x$ does not exist. However, the sequence has a very simple pattern:

$$1,-1,1,-1,1\ldots$$

and clearly diverges. ♣

Example 6.12: Convergence of Exponential Terms

Determine whether $\{(-1/2)^n\}_{n=0}^{\infty}$ *converges or diverges. If it converges, compute the limit.*

Solution. We consider the absolute value of the sequence $\{|-1/2|^n\}_{n=0}^{\infty}=\{(1/2)^n\}_{n=0}^{\infty}$ and define

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

Then

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

so by the Absolute Value Sequence Theorem the sequence converges to 0. ♣

Example 6.13: Using the Squeeze Theorem for Sequences

Determine whether $\left\{\frac{\sin n}{\sqrt{n}}\right\}_{n=1}^{\infty}$ *converges or diverges. If it converges, compute the limit.*

Solution. Since $|\sin n| \leq 1$ for all n, we have that

$$0 \leq \left|\frac{\sin n}{\sqrt{n}}\right| \leq \frac{1}{\sqrt{n}}.$$

Let $a_n = 0$ and $c_n = \dfrac{1}{\sqrt{n}}$, then

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} 0 = 0 \quad \text{and} \quad \lim_{n\to\infty} c_n = \lim_{n\to\infty} \frac{1}{\sqrt{n}} = 0.$$

Applying the Squeeze Theorem, we have

$$\lim_{n\to\infty} \frac{\sin n}{\sqrt{n}} = 0,$$

and the sequence converges to 0. ♣

Example 6.14: Geometric Sequence

Let r be a fixed real number. Determine when $\{r^n\}_{n=0}^{\infty}$ coverges.

Solution. A particularly common and useful sequence is $\{r^n\}_{n=0}^{\infty}$, for various values of r. Some are quite easy to understand:

- If $r = 1$ the sequence converges to 1 since every term is 1, and likewise if $r = 0$ the sequence converges to 0.
- If $r = -1$ this is the sequence of Example 6.11 and diverges.
- If $r > 1$ or $r < -1$ the terms r^n get large without limit, so the sequence diverges.
- If $0 < r < 1$ then the sequence converges to 0.
- If $-1 < r < 0$ then $|r^n| = |r|^n$ and $0 < |r| < 1$, so the sequence $\{|r|^n\}_{n=0}^{\infty}$ converges to 0, so also $\{r^n\}_{n=0}^{\infty}$ converges to 0.

In summary, $\{r^n\}$ converges precisely when $-1 < r \leq 1$ in which case

$$\lim_{n\to\infty} r^n = \begin{cases} 0 & \text{if } -1 < r < 1 \\ 1 & \text{if } r = 1 \end{cases}$$

♣

Note: Sequences of this form, or the more general form $\{kr^n\}_{n=0}^{\infty}$, are called **geometric sequences** or **geometric progressions**. They are encountered in a large variety of mathematical and real-world applications.

Definition 6.15: Geometric Sequence

A **geometric sequence** *is of the form*

$$\{a_n\}_{n=0}^{\infty} = \{kr^n\}_{n=0}^{\infty},$$

where k is the first term and $r = \dfrac{a_{n+1}}{a_n}$ for $n \geq 0$ is the **common ratio**.

Sometimes we will not be able to determine the limit of a sequence, but we still would like to know whether it converges. In some cases we can determine this even without being able to compute the limit. For this, we introduce two basic properties of sequences, namely **monotonicity** and **boundedness**.

Definition 6.16: Monotonic Sequence

A sequence $\{a_n\}$ wih the following properties is called **monotonic***:*

1. *If $a_n < a_{n+1}$ for all n, then the sequence is* **increasing** *or* **strictly increasing**.
2. *If $a_n \leq a_{n+1}$ for all n, then the sequence is* **non-decreasing**.
3. *If $a_n > a_{n+1}$ for all n, then the sequence is* **decreasing** *or* **strictly decreasing**.
4. *If $a_n \geq a_{n+1}$ then the sequence is* **non-increasing**.

Note: Unfortunately, some authors refer to a *non-decreasing* sequence as *increasing* or, similarly, to a *non-increasing* sequence as *decreasing*, which is not correct in the strict sense. For example, the sequence $\{a_1 = 0,\ a_2 = 0,\ a_3 = 1\}$ is not increasing (or strictly increasing), since $a_1 = a_2 = 0$; however, it is non-decreasing because $a_n \leq a_{n+1}$ for all n.

Example 6.17: Monotonic or Not

Determine the monotonicity of the following sequences:

(a) $\left\{\dfrac{2^i - 1}{2^i}\right\}_{i=1}^{\infty} = \dfrac{1}{2}, \dfrac{3}{4}, \dfrac{7}{8}, \dfrac{15}{16}, \ldots$ *(b)* $\left\{\dfrac{n+1}{n}\right\}_{n=1}^{\infty} = \dfrac{2}{1}, \dfrac{3}{2}, \dfrac{4}{3}, \dfrac{5}{4}, \ldots$

Solution.

(a) We investigate monotonicity by comparing the i-th and $(i+1)$-th terms:

Left side	Right side
$\dfrac{2^i - 1}{2^i}$	$\dfrac{2^{i+1} - 1}{2^{i+1}}$
$2(2^i - 1)$	$2^{i+1} - 1$
$2^{i+1} - 2$	$2^{i+1} - 1$
-2	-1

From the last statement, we observe that $-2 < -1$, and so

$$\frac{2^i - 1}{2^i} < \frac{2^{i+1} - 1}{2^{i+1}}$$

for all i. Hence, the sequence is strictly increasing.

(b) We need to compare the n-th and $n+1$-th terms:

Left side	Right side
$\frac{n+1}{n}$	$\frac{(n+1)+1}{(n+1)}$
$\frac{n+1}{n}$	$\frac{n+2}{n+1}$
$(n+1)(n+1)$	$n(n+2)$
n^2+2n+1	n^2+2n
1	0

From the last statement, we observe that $1 > 0$ and so

$$\frac{n+1}{n} > \frac{(n+1)+1}{(n+1)}$$

for all n. Hence, the sequence is strictly decreasing.

♣

Definition 6.18: Bounded Sequences

A sequence $\{a_n\}$ is **bounded above** *if there is some number N such that $a_n \leq N$ for every n, and* **bounded below** *if there is some number N such that $a_n \geq N$ for every n. If a sequence is bounded above and bounded below it is* **bounded**.

Note: If a sequence $\{a_n\}_{n=0}^{\infty}$ is increasing or non-decreasing it is bounded below (by a_0), and if it is decreasing or non-increasing it is bounded above (by a_0).

Finally, with all this new terminology we can state an important theorem concerning the convergence of a monotonic and increasing sequence.

Theorem 6.19: Bounded Monotonic Sequence

If a sequence is bounded and monotonic then it converges.

We will not prove this, but the proof appears in many calculus books. It is not hard to believe: suppose that a sequence is increasing and bounded, so each term is larger than the one before, yet never larger than some fixed value N. The terms must then get closer and closer to some value between a_0 and N. It need not be N, since N may be a "too-generous" upper bound; the limit will be the smallest number that is above all of the terms a_i.

Example 6.20: Converging or Diverging

Determine whether $\left\{\dfrac{2^i-1}{2^i}\right\}_{i=1}^{\infty}$ *converges.*

Solution. For every $i \geq 1$ we have

$$0 < \frac{(2^i-1)}{2^i} < 1,$$

so the sequence is bounded, and we have already observed that it is strictly increasing in Example 6.17. Therefore, the sequence converges according to the Bounded Monotonic Sequence Theorem. ♣

Note: We don't actually need to know that a sequence is monotonic to apply this theorem – it is enough to know that the sequence is "eventually" monotonic, that is, that at some point it becomes increasing or decreasing. For example, the sequence 10, 9, 8, 15, 3, 21, 4, $3/4$, $7/8$, $15/16$, $31/32, \ldots$ is not increasing, because among the first few terms it is not. But starting with the term $3/4$ it is increasing, so the theorem tells us that the sequence $3/4, 7/8, 15/16, 31/32, \ldots$ converges. Since convergence depends only on what happens as n gets large, adding a few terms at the beginning can't turn a convergent sequence into a divergent one.

Example 6.21: Demonstrating Convergence

Show that $\{n^{1/n}\}$ *converges.*

Solution. We first show that this sequence is decreasing by showing that

$$n^{1/n} > (n+1)^{1/(n+1)}.$$

Consider the real function $f(x) = x^{1/x}$ when $x \geq 1$. We can compute the derivative,

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

and note that when $x \geq 3$ this is negative. Since the function has negative slope,

$$n^{1/n} > (n+1)^{1/(n+1)}$$

when $n \geq 3$.

Since all terms of the sequence are positive, the sequence is decreasing and bounded when $n \geq 3$, and so the sequence converges. (As it happens, we can compute the limit in this case, but we know it converges even without knowing the limit; see Exercise 6.1.2.) ♣

Note: When we need to determine whether a sequence $\{a_n\}$ is monotonic, it is often useful to investigate the ratio of successive terms $\dfrac{a_{n+1}}{a_n}$, as shown in the following example.

Example 6.22: Demonstrating Convergence

Show that $\left\{\dfrac{n!}{n^n}\right\}$ *converges.*

Solution. If we look at the ratio of successive terms we see that:

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)!}{(n+1)^{n+1}}\frac{n^n}{n!} = \frac{(n+1)!}{n!}\frac{n^n}{(n+1)^{n+1}} = \frac{n+1}{n+1}\left(\frac{n}{n+1}\right)^n = \left(\frac{n}{n+1}\right)^n < 1.$$

Therefore $a_{n+1} < a_n$, and so the sequence is decreasing. Since all terms are positive, it is also bounded, and so it must converge. (Again it is possible to compute the limit; see Exercise 6.1.3.) ♣

Exercises for Section 6.1

Exercise 6.1.1 *Graph the first 5 terms of the following sequences.*

(a) $\{n^2 - n\}_{n=0}^{\infty}$

(b) $\left\{\frac{\sqrt{n}}{\sqrt{n}+1}\right\}_{n=0}^{\infty}$

(c) $\left\{\frac{e^n}{n^2}\right\}_{n=1}^{\infty}$

(d) $\left\{\frac{n}{3^n}\right\}_{n=1}^{\infty}$

Exercise 6.1.2 *Compute* $\lim\limits_{x\to\infty} x^{1/x}$.

Exercise 6.1.3 *Use the Squeeze Theorem to show that* $\lim\limits_{n\to\infty}\frac{n!}{n^n} = 0$.

Exercise 6.1.4 *Determine whether the following sequences converge or diverge. In the case of convergence, compute the limit.*

(a) $\{n^2 - n\}_{n=0}^{\infty}$

(b) $\left\{\frac{\sqrt{n}}{\sqrt{n}+1}\right\}_{n=0}^{\infty}$

(c) $\left\{\frac{e^n}{n^2}\right\}_{n=1}^{\infty}$

(d) $\left\{\frac{n}{3^n}\right\}_{n=1}^{\infty}$

(e) $\left\{\sqrt{n+47} - \sqrt{n}\right\}_{n=0}^{\infty}$

(f) $\left\{\frac{n^2+1}{(n+1)^2}\right\}_{n=0}^{\infty}$

(g) $\left\{\frac{n+47}{\sqrt{n^2+3n}}\right\}_{n=1}^{\infty}$

(h) $\left\{\frac{2^n}{n!}\right\}_{n=0}^{\infty}$

6.2 Series

While much more can be said about sequences, we now turn to our principal interest, **series**. Recall that a series, roughly speaking, is the sum of a sequence: If $\{a_n\}_{n=0}^{\infty}$ is a sequence then we can construct a series by adding up all the terms in the sequence:

$$\sum_{n=0}^{\infty} a_n = a_0 + a_1 + a_2 + \cdots$$

Definition 6.23: Infinite Series

Given an infinite sequence $\{a_n\}_{n=0}^{\infty}$, the sum

$$\sum_{n=0}^{\infty} a_n = a_0 + a_1 + a_2 + \ldots$$

is called an **infinite series**.

Associated with a series is a second sequence, called the **sequence of partial sums** $\{s_n\}_{n=0}^{\infty}$, where the n-th partial sum s_n always terminates with the n-th term a_n of the sequence:

$$s_0 = a_0, \quad s_1 = a_0 + a_1, \quad s_2 = a_0 + a_1 + a_2, \quad \ldots$$

Definition 6.24: Sequence of Partial Sums

Given an infinite series $\sum_{n=0}^{\infty} a_n$, the **sequence of partial sums** *is $\{s_n\}_{n=0}^{\infty}$, where the* **n-th partial sum** *is defined as*

$$s_n = \sum_{i=0}^{n} a_i = a_0 + a_1 + a_2 + \cdots + a_n.$$

Again, we are interested in the behaviour of the series, and whether its terms sum to a finite number or not. We therefore define **convergence** and **divergence** for a series.

Definition 6.25: Convergence and Divergence of a Series

Given an infinite series $\sum_{n=0}^{\infty} a_n$, the series **converges** *if the associated sequence of partial sums converges; otherwise, the series* **diverges**.

Definition 6.26: Geometric Series

If $\{kx^n\}_{n=0}^{\infty}$ is a geometric sequence, then the associated series $\sum_{i=0}^{\infty} kx^i$ is called a **geometric series**.

Note: An infinite series $\sum_{n=0}^{\infty} a_n$ can also be denoted by $\sum_{n} a_n$ or $\sum a_n$, where it is understood that n is an integer that runs over all admissible values of the index. However, if we are interested in the actual sum of the series, then it is pertinent to identify the lower bound of the series.

Theorem 6.27: Geometric Series Convergence

Given a geometric series $\sum_{n=0}^{\infty} kx^n$ *with* $k \neq 0$*:*

1. *If* $|x| < 1$*, the series converges*
$$\sum_{n=0}^{\infty} kx^n = \frac{k}{1-x}.$$
2. *If* $|x| \geq 1$*, the series diverges.*

Proof.

A typical partial sum is
$$s_n = k + kx + kx^2 + kx^3 + \cdots + kx^n = k(1 + x + x^2 + x^3 + \cdots + x^n).$$

We note that for $x \neq 1$,
$$\begin{aligned} s_n(1-x) &= k(1 + x + x^2 + x^3 + \cdots + x^n)(1-x) \\ &= k(1 + x + x^2 + x^3 + \cdots + x^n)1 - k(1 + x + x^2 + x^3 + \cdots + x^{n-1} + x^n)x \\ &= k(1 + x + x^2 + x^3 + \cdots + x^n - x - x^2 - x^3 - \cdots - x^n - x^{n+1}) \\ &= k(1 - x^{n+1}) \end{aligned}$$

so
$$\begin{aligned} s_n(1-x) &= k(1 - x^{n+1}) \\ s_n &= k\frac{1 - x^{n+1}}{1-x}. \end{aligned}$$

1. If $|x| < 1$, $\lim_{n\to\infty} x^n = 0$ so
$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} k\frac{1 - x^{n+1}}{1-x} = k\frac{1}{1-x}.$$
Thus, when $|x| < 1$ the geometric series converges to $k/(1-x)$.
2. (a) If $|x| > 1$, then
$$\lim_{n\to\infty} x^n = \infty$$
and so the series diverges.

   (b) If $|x| = 1$, then
$$|s_n| = \sum_{i=0}^{n} 1 = n + 1.$$
Therefore,
$$\lim_{n\to\infty} |s_n| = \lim_{n\to\infty} (n+1) = \infty$$
and so the series diverges.

♣

Example 6.28: Summing a Geometric Series

Given the series $\sum_{n=0}^{\infty} \frac{1}{2^n}$*, determine the following:*

(a) The partial sum s_n*.*

(b) The sum of the series.

Solution. We recognize that the series is geometric with $k = 1$ and $x = 1/2$.

(a) $s_n = \dfrac{1-(1/2)^{n+1}}{1-1/2} = \dfrac{2^{n+1}-1}{2^n} = 2 - \dfrac{1}{2^n}$

(b) $\displaystyle\sum_{n=0}^{\infty} \frac{1}{2^n} = \frac{1}{1-1/2} = 2.$

♣

We began the chapter with the series

$$\sum_{n=1}^{\infty} \frac{1}{2^n},$$

namely, the geometric series without the first term 1. Each partial sum of this series is 1 less than the corresponding partial sum for the geometric series, so of course the limit is also one less than the value of the geometric series, that is,

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = 1.$$

It is not hard to see that the following theorem follows from Theorem 6.6.

Theorem 6.29: Series are Linear

Suppose that $\sum a_n$ *and* $\sum b_n$ *are convergent series, and* c *is a constant. Then*

1. $\sum c a_n$ *is convergent and* $\sum c a_n = c \sum a_n$
2. $\sum (a_n + b_n)$ *is convergent and* $\sum (a_n + b_n) = \sum a_n + \sum b_n$*.*

Note:

1. When c is non-zero, the converse of the first part of this theorem is also true. That is, if $\sum c a_n$ is convergent, then $\sum a_n$ is also convergent; if $\sum c a_n$ converges then $\frac{1}{c} \sum c a_n$ must converge.

2. On the other hand, the converse of the second part of the theorem is not true. For example, if $a_n = 1$ and $b_n = -1$, then $\sum a_n + \sum b_n = \sum 0 = 0$ converges, but each of $\sum a_n$ and $\sum b_n$ diverges.

In general, the sequence of partial sums s_n is harder to understand and analyze than the sequence of terms a_n, and it is difficult to determine whether series converge and if so to what. The following result will let us deal with some simple cases easily.

Theorem 6.30: Divergence Test

If $\sum a_n$ *converges then* $\lim_{n\to\infty} a_n = 0$.

Proof. Since $\sum a_n$ converges, $\lim_{n\to\infty} s_n = L$ and $\lim_{n\to\infty} s_{n-1} = L$, because this really says the same thing but "renumbers" the terms. By Theorem 6.6,

$$\lim_{n\to\infty}(s_n - s_{n-1}) = \lim_{n\to\infty} s_n - \lim_{n\to\infty} s_{n-1} = L - L = 0.$$

But

$$s_n - s_{n-1} = (a_0 + a_1 + a_2 + \cdots + a_n) - (a_0 + a_1 + a_2 + \cdots + a_{n-1}) = a_n,$$

so as desired $\lim_{n\to\infty} a_n = 0$. ♣

Note:

1. This theorem presents an easy Divergence Test when we use the contrapositive form: Given a series $\sum a_n$, if the limit $\lim_{n\to\infty} a_n$ does not exist or has a value other than zero, the series diverges. This result is captured in the next theorem called the ***n*-th Term Test**. Often, this theorem is referred to as the *Divergence Test*.

2. Note well that the converse is *not* true: If $\lim_{n\to\infty} a_n = 0$ then the series does not necessarily converge. One example is the so-called **harmonic series**, as will be shown in Example 6.33.

Theorem 6.31: *n*-th Term Test

If $\lim_{n\to\infty} a_n \neq 0$ *or if the limit does not exist, then* $\sum a_n$ *diverges.*

Proof. Consider the statement of the theorem in contrapositive form:

$$\text{If } \sum_{n=1}^{\infty} a_n \text{ converges, then } \lim_{n\to\infty} a_n = 0.$$

If s_n are the partial sums of the series, then the assumption that the series converges gives us

$$\lim_{n\to\infty} s_n = s$$

for some number s. Then

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty}(s_n - s_{n-1}) = \lim_{n\to\infty} s_n - \lim_{n\to\infty} s_{n-1} = s - s = 0.$$

♣

Example 6.32: Demonstrating Divergence

Show that $\displaystyle\sum_{n=1}^{\infty} \frac{n}{n+1}$ *diverges.*

Solution. We compute the limit:

$$\lim_{n\to\infty} \frac{n}{n+1} = 1 \neq 0.$$

Looking at the first few terms perhaps makes it clear that the series has no chance of converging:

$$\frac{1}{2}+\frac{2}{3}+\frac{3}{4}+\frac{4}{5}+\cdots$$

will just get larger and larger; indeed, after a bit longer the series starts to look very much like $\cdots+1+1+1+1+\cdots$, and of course if we add up enough 1's we can make the sum as large as we desire. ♣

Example 6.33: Harmonic Series

Show that $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$ *diverges.*

Solution. Here the theorem does not apply: $\lim\limits_{n\to\infty} 1/n = 0$, so it looks like perhaps the series converges. Indeed, if you have the fortitude (or the software) to add up the first 1000 terms you will find that

$$\sum_{n=1}^{1000} \frac{1}{n} \approx 7.49,$$

so it might be reasonable to speculate that the series converges to something in the neighborhood of 10. But in fact the partial sums do go to infinity; they just get big very, very slowly. Consider the following:

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4} > 1+\frac{1}{2}+\frac{1}{4}+\frac{1}{4} = 1+\frac{1}{2}+\frac{1}{2}$$
$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8} > 1+\frac{1}{2}+\frac{1}{4}+\frac{1}{4}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8} = 1+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}$$
$$1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{16} > 1+\frac{1}{2}+\frac{1}{4}+\frac{1}{4}+\frac{1}{8}+\cdots+\frac{1}{8}+\frac{1}{16}+\cdots+\frac{1}{16} = 1+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}$$

and so on. By swallowing up more and more terms we can always manage to add at least another $1/2$ to the sum, and by adding enough of these we can make the partial sums as big as we like. In fact, it's not hard to see from this pattern that

$$1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{2^n} > 1+\frac{n}{2},$$

so to make sure the sum is over 100, for example, we'd add up terms until we get to around $1/2^{198}$, that is, about $4\cdot 10^{59}$ terms. ♣

Definition 6.34: Harmonic Series

A series of the form

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \ldots$$

is called a **harmonic series**.

Note: We will often make use of the fact that the first few (e.g. any finite number of) terms in a series are irrelevant when determining whether it will converge. In other words, $\sum_{n=0}^{\infty} a_n$ converges if and only if $\sum_{n=N}^{\infty} a_n$ converges for some $N \geq 1$.

Exercises for Section 6.2

Exercise 6.2.1 *Explain why the following series diverge.*

(a) $\displaystyle\sum_{n=1}^{\infty} \frac{n^2}{2n^2+1}$

(b) $\displaystyle\sum_{n=1}^{\infty} \frac{5}{2^{1/n}+14}$

(c) $\displaystyle\sum_{n=1}^{\infty} \frac{3}{n}$

Exercise 6.2.2 *Compute the sum of the following series, if it converges.*

(a) $\displaystyle\sum_{n=0}^{\infty} \left(\frac{4}{(-3)^n} - \frac{3}{3^n}\right)$

(b) $\displaystyle\sum_{n=0}^{\infty} \left(\frac{3}{2^n} + \frac{4}{5^n}\right)$

(c) $\displaystyle\sum_{n=0}^{\infty} \frac{4^{n+1}}{5^n}$

(d) $\displaystyle\sum_{n=0}^{\infty} \frac{3^{n+1}}{7^{n+1}}$

(e) $\displaystyle\sum_{n=1}^{\infty} \left(\frac{3}{5}\right)^n$

(f) $\displaystyle\sum_{n=1}^{\infty} \frac{3^n}{5^{n+1}}$

6.3 Integral Test

It is generally quite difficult, often impossible, to determine the value of a series exactly. In many cases it is possible at least to determine whether or not the series converges, and so we will spend most of our time on this problem.

If all of the terms a_n in a series are non-negative, then clearly the sequence of partial sums s_n is non-decreasing. This means that if we can show that the sequence of partial sums is bounded, the series must converge. Many useful and interesting series have this property, and they are among the easiest to understand. Let's look at an example.

Example 6.35: Exploring Convergence Using an Integral

Show that $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2}$ *converges.*

Solution. The terms $1/n^2$ are positive and decreasing, and since $\lim\limits_{x\to\infty} 1/x^2 = 0$, the terms $1/n^2$ approach zero. This means that the Divergence Test does not provide any information and we must find a different method to deal with this series. We seek an upper bound for all the partial sums, that is, we want to find a number N so that $s_n \leq N$ for every n. The upper bound is provided courtesy of integration, and is illustrated in Figure 6.2.

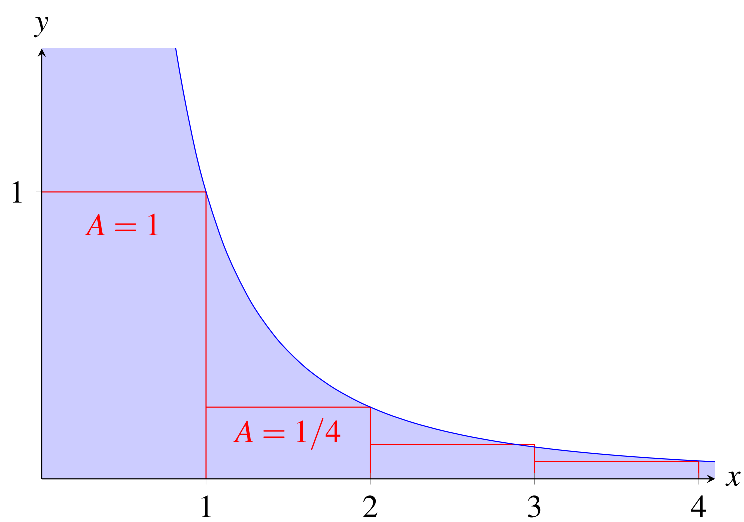


Figure 6.2: Graph of $y = 1/x^2$ with rectangles.

The figure shows the graph of $y = 1/x^2$ together with some rectangles that lie completely below the curve and that all have base length one. Because the heights of the rectangles are determined by the height of the curve, the areas of the rectangles are $1/1^2$, $1/2^2$, $1/3^2$, and so on—in other words, exactly the terms of the series. The partial sum s_n is simply the sum of the areas of the first n rectangles. Because the rectangles all lie between the curve and the x-axis, any sum of rectangle areas is less than the corresponding area under the curve, and so of course any sum of rectangle areas is less than the area under the entire curve. Unfortunately, because of the asymptote at $x = 0$, the integral $\int_0^\infty \frac{1}{x^2}$ is infinite, but we can deal with this by separating the first term from the series and integrating from 1:

$$s_n = \sum_{i=1}^{n} \frac{1}{i^2} = 1 + \sum_{i=2}^{n} \frac{1}{i^2} < 1 + \int_1^n \frac{1}{x^2}\,dx < 1 + \int_1^\infty \frac{1}{x^2}\,dx = 1 + 1 = 2,$$

recalling that we computed this improper integral in Section 2.7. Since the sequence of partial sums s_n is increasing and bounded above by 2, we know that $\lim_{n\to\infty} s_n = L < 2$, and so the series converges to some number less than 2. In fact, it is possible, though difficult, to show that $L = \pi^2/6 \approx 1.6$. ♣

Example 6.36: Exploring Divergence Using an Integral

Why can the integral technique from Example 6.35 not be used with the series

$$\sum_{n=1}^{\infty} \frac{1}{n}?$$

Solution. We already know that $\sum 1/n$ diverges. What goes wrong if we try to apply the integral technique to it? Here's the calculation:

$$s_n = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} < 1 + \int_1^n \frac{1}{x}\,dx < 1 + \int_1^\infty \frac{1}{x}\,dx = 1 + \infty.$$

The problem is that the improper integral doesn't converge. Note that this does *not* prove that $\sum 1/n$ diverges, just that this particular technique fails to prove that it converges. ♣

A slight modification, however, allows us to prove in a second way that $\sum 1/n$ diverges.

Example 6.37: Alternate Method for Divergence of Harmonic Series

Use the idea of areas from Example 6.35 to show that the series

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

diverges.

Solution. Consider a slightly altered version of Figure 6.2, shown in Figure 6.3.

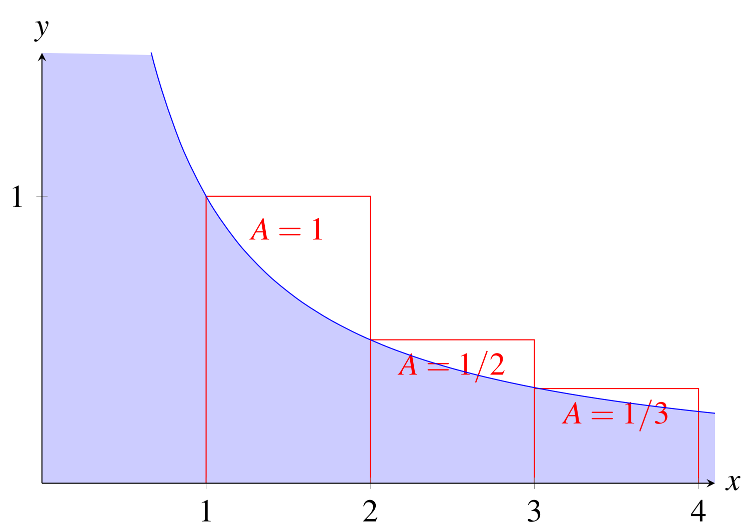


Figure 6.3: Graph of $y = 1/x$ with rectangles.

This time the rectangles are above the curve, that is, each rectangle completely contains the corresponding area under the curve. This means that

$$s_n = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} > \int_1^{n+1} \frac{1}{x}\,dx = \ln x \Big|_1^{n+1} = \ln(n+1).$$

As n gets bigger, $\ln(n+1)$ goes to infinity, so the sequence of partial sums s_n must also go to infinity, so the harmonic series diverges.

The key fact in this example is that

$$\lim_{n\to\infty} \int_1^{n+1} \frac{1}{x}\,dx = \int_1^{\infty} \frac{1}{x}\,dx = \infty.$$

♣

So these two examples taken together indicate that we can prove that a series converges or prove that it diverges with a single calculation of an improper integral. This is known as the **Integral Test**, which we state as a theorem.

Theorem 6.38: Integral Test

Suppose that f is a continuous, positive, and decreasing function of x on the infinite interval $[1,\infty)$ and that $a_n = f(n)$. Then

$$\sum_{n=1}^{\infty} a_n \quad \text{and} \quad \int_1^{\infty} f(x)\,dx$$

either both converge or both diverge.

Note: The lower bound in the Integral Test is arbitrary. We could have chosen any positive integer N as the lower bound, since – as mentioned before – the first few (e.g. any finite number of) terms in a series are irrelevant when determining whether it will converge.

The two examples we have seen are called p**-series**. A p-series is any series of the form $\sum 1/n^p$.

Definition 6.39: p-Series

A series of the form $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^p} = 1 + \frac{1}{2^p} + \frac{1}{3^p} + \cdots + \frac{1}{n^p} + \ldots$, *where p is a constant, is called a* **p-series**.

Theorem 6.40: p-Series Test

Given the p-series

$$\sum_{n=1}^{\infty} \frac{1}{n^p} = 1 + \frac{1}{2^p} + \frac{1}{3^p} + \cdots + \frac{1}{n^p} + \cdots :$$

1. *If $p > 1$, the series converges.*
2. *If $p \leq 1$, the series diverges.*

Proof. We split the proof into three cases:

$$\int_1^\infty \frac{1}{x^p}\,dx = \lim_{R\to\infty} \frac{x^{1-p}}{1-p}\bigg|_1^R = \lim_{R\to\infty} \frac{R^{1-p}}{1-p} - \frac{1}{1-p}.$$

1. If $p > 1$ then $1-p < 0$ and $\lim_{R\to\infty} R^{1-p} = 0$, so the integral converges.
2. If $p < 1$ then $1-p > 0$ and $\lim_{R\to\infty} R^{1-p} = \infty$, so the integral diverges.
3. If $p = 1$, we have the harmonic series, which we have already shown in two ways to be divergent.

♣

Example 6.41: p-Series Power of Three

Show that $\sum_{n=1}^{\infty} \frac{1}{n^3}$ *converges.*

Solution. We could of course use the Integral Test, but now that we have the theorem we may simply note that this is a p-series with $p > 1$. ♣

Example 6.42: p-Series Power of Four

Show that $\sum_{n=1}^{\infty} \frac{5}{n^4}$ *converges.*

Solution. We know that if $\sum_{n=1}^{\infty} 1/n^4$ converges then $\sum_{n=1}^{\infty} 5/n^4$ also converges, by Theorem 6.29. Since $p = 4 > 1$ we have that $\sum_{n=1}^{\infty} 1/n^4$ is a convergent p-series, and so $\sum_{n=1}^{\infty} 5/n^4$ converges also. ♣

Example 6.43: p-Series Square Root

Show that $\sum_{n=1}^{\infty} \frac{5}{\sqrt{n}}$ *diverges.*

Solution. This also follows from Theorem 6.29: Since $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$ is a p-series with $p = 1/2 < 1$, it diverges, and so does $\sum_{n=1}^{\infty} \frac{5}{\sqrt{n}}$. ♣

Note: Since it is typically difficult to compute the value of a series exactly, a good approximation is frequently required. In a real sense, a good approximation is only as good as we know it is, that is, while an approximation may in fact be good, it is only valuable in practice if we can guarantee its accuracy to some degree. This guarantee is usually easy to come by for series with decreasing positive terms.

Example 6.44: Approximating a p-Series

Approximate $\sum_{n=1}^{\infty} 1/n^2$ to within 0.01.

Solution. Referring to Figure 6.2, if we approximate the sum by $\sum_{n=1}^{N} 1/n^2$, the size of the error we make is the total area of the remaining rectangles, all of which lie under the curve $1/x^2$ from $x = N$ to infinity. So we know the true value of the series is larger than the approximation, and no bigger than the approximation plus the area under the curve from N to infinity. Roughly, then, we need to find N so that

$$\int_N^\infty \frac{1}{x^2}\,dx < 1/100.$$

We can compute the integral:

$$\int_N^\infty \frac{1}{x^2}\,dx = \frac{1}{N},$$

so if we choose $N = 100$ the error will be less than 0.01. Adding up the first 100 terms gives approximately 1.634983900. In fact, we can do a bit better. Since we know that the correct value is between our approximation and our approximation plus the error (not minus), we can cut our error bound in half by taking the value midway between these two values. If we take $N = 50$, we get a sum of 1.6251327 with an error of at most 0.02, so the correct value is between 1.6251327 and 1.6451327, and therefore the value halfway between these, 1.6351327, is within 0.01 of the correct value. We have mentioned that the true value of this series can be shown to be $\pi^2/6 \approx 1.644934068$ which is 0.0098 more than our approximation, and so (just barely) within the required error. Frequently approximations will be even better than the "guaranteed" accuracy, but not always, as this example demonstrates. ♣

Exercises for Section 6.3

Exercise 6.3.1 *Determine whether each series converges or diverges.*

(a) $\sum_{n=1}^{\infty} \frac{1}{n^{\pi/4}}$

(b) $\sum_{n=1}^{\infty} \frac{n}{n^2+1}$

(c) $\sum_{n=1}^{\infty} \frac{\ln n}{n^2}$

(d) $\sum_{n=1}^{\infty} \frac{1}{n^2+1}$

(e) $\sum_{n=1}^{\infty} \frac{1}{e^n}$

(f) $\sum_{n=1}^{\infty} \frac{n}{e^n}$

(g) $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n \ln n}$

(h) $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^2}$

Exercise 6.3.2 *Find an N so that each series is approximated to within the given error.*

(a) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^4}$, 0.005

(b) $\displaystyle\sum_{n=0}^{\infty} \frac{1}{e^n}$, 10^{-4}

(c) $\displaystyle\sum_{n=1}^{\infty} \frac{\ln n}{n^2}$, 0.005

(d) $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^2}$, 0.005

6.4 Alternating Series

Next we consider series with both positive and negative terms, but in a regular pattern: they alternate signs. For example:

- $\displaystyle\sum_{n=1}^{\infty} (-1)^{n-1} \frac{2}{n+1} = 1 - \frac{2}{3} + \frac{2}{4} - \frac{2}{5} + \dots$
- $\displaystyle\sum_{n=1}^{\infty} (-1)^{n} \frac{2}{n+1} = -1 + \frac{2}{3} - \frac{2}{4} + \frac{2}{5} - \dots$
- $\displaystyle\sum_{n=0}^{\infty} (-1)^{n} 2^n = 1 - 2 + 4 - 8 + \dots$
- $\displaystyle\sum_{n=4}^{\infty} (-1)^{n-1} \frac{n}{n+2} = -\frac{4}{6} + \frac{5}{7} - \frac{6}{8} + \frac{7}{9} - \dots$

Definition 6.45: Alternating Series

An alternating series has the form

$$\sum (-1)^n a_n$$

where a_n are all positive and the first index is arbitrary.

Note: An alternating series can start with a positive or negative term, i.e. the first index can be any non-negative integer.

A well-known example of an alternating series is the **alternating harmonic series**:

Definition 6.46: Alternating Harmonic Series

A series of the form

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n-1}}{n} + \ldots$$

is called an **alternating harmonic series**.

In the alternating harmonic series the magnitude of the terms decrease, that is, $|a_n|$ forms a decreasing sequence, although this is not required in an alternating series. Recall that for a series with positive terms, if the limit of the terms is not zero, the series cannot converge; but even if the limit of the terms is zero, the series still may not converge. It turns out that for alternating series, the series converges exactly when the limit of the terms is zero. In Figure 6.4, we illustrate what happens to the partial sums of the alternating harmonic series. Because the sizes of the terms a_n are decreasing, the odd partial sums s_1, s_3, s_5, and so on, form a decreasing sequence that is bounded below by s_2, so this sequence must converge. Likewise, the even partial sums s_2, s_4, s_6, and so on, form an increasing sequence that is bounded above by s_1, so this sequence also converges. Since all the even numbered partial sums are less than all the odd numbered ones, and since the "jumps" (that is, the a_i terms) are getting smaller and smaller, the two sequences must converge to the same value, meaning the entire sequence of partial sums $s_1, s_2, s_3, \ldots$ converges as well.

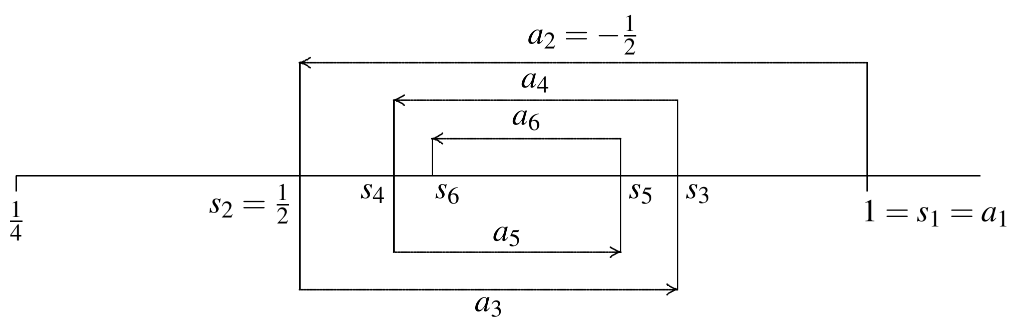


Figure 6.4: The alternating harmonic series.

The same argument works for any alternating sequence with terms that decrease in absolute value. The Alternating Series Test is worth calling a theorem.

Theorem 6.47: Alternating Series Test

Suppose that $\{a_n\}_{n=1}^{\infty}$ *is a decreasing sequence of positive numbers and* $\lim\limits_{n\to\infty} a_n = 0$. *Then the alternating series* $\sum\limits_{n=1}^{\infty}(-1)^{n-1}a_n$ *converges.*

Proof. The odd-numbered partial sums, $s_1, s_3, s_5, \ldots, s_{2k+1}, \ldots$, form a decreasing sequence, because $s_{2k+3} = s_{2k+1} - a_{2k+2} + a_{2k+3} \leq s_{2k+1}$, since $a_{2k+2} \geq a_{2k+3}$. This sequence is bounded below by s_2, so it must converge, to some value L. Likewise, the partial sums $s_2, s_4, s_6, \ldots, s_{2k}, \ldots$, form an increasing sequence that is bounded above by s_1, so this sequence also converges, to some value M. Since $\lim\limits_{n\to\infty} a_n = 0$ and

$s_{2k+1} = s_{2k} + a_{2k+1}$,

$$L = \lim_{k\to\infty} s_{2k+1} = \lim_{k\to\infty}(s_{2k} + a_{2k+1}) = \lim_{k\to\infty} s_{2k} + \lim_{k\to\infty} a_{2k+1} = M + 0 = M,$$

so $L = M$; the two sequences of partial sums converge to the same limit, and this means the entire sequence of partial sums also converges to L. ♣

Another useful fact is implicit in this discussion. Suppose that

$$L = \sum_{n=1}^{\infty}(-1)^{n-1}a_n$$

and that we approximate L by a finite part of this sum, say

$$L \approx \sum_{n=1}^{N}(-1)^{n-1}a_n.$$

Because the terms are decreasing in size, we know that the true value of L must be between this approximation and the next one, that is, between

$$\sum_{n=1}^{N}(-1)^{n-1}a_n \quad \text{and} \quad \sum_{n=1}^{N+1}(-1)^{n-1}a_n.$$

Depending on whether N is odd or even, the second will be smaller or larger than the first.

Example 6.48: Approximating a Series

Approximate the sum of the alternating harmonic series to within 0.05.

Solution. We need to go to the point at which the next term to be added or subtracted is $1/10$. Adding up the first nine and the first ten terms we get approximately 0.746 and 0.646. These are $1/10$ apart, so the value halfway between them, 0.696, is within 0.05 of the correct value. ♣

Note: We have considered alternating series with first index 1, and in which the first term is positive, but a little thought shows this is not crucial. The same test applies to any similar series, such as $\sum_{n=0}^{\infty}(-1)^n a_n$, $\sum_{n=1}^{\infty}(-1)^n a_n$, $\sum_{n=17}^{\infty}(-1)^n a_n$, etc.

Exercises for Section 6.4

Exercise 6.4.1 *Determine whether the following series converge or diverge.*

(a) $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2n+5}$

(b) $\displaystyle\sum_{n=4}^{\infty} \frac{(-1)^{n-1}}{\sqrt{n-3}}$

(c) $\displaystyle\sum_{n=1}^{\infty} (-1)^{n-1} \frac{n}{3n-2}$

(d) $\displaystyle\sum_{n=1}^{\infty} (-1)^{n-1} \frac{\ln n}{n}$

Exercise 6.4.2 *Approximate each series to within the given error.*

(a) $\displaystyle\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n^3}$, *0.005*

(b) $\displaystyle\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n^4}$, *0.005*

6.5 Comparison Test

As we begin to compile a list of convergent and divergent series, new ones can sometimes be analyzed by comparing them to ones that we already understand.

Example 6.49: Convergence by Comparison

Does $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n^2 \ln n}$ *converge?*

Solution. The obvious first approach, based on what we know, is the Integral Test. Unfortunately, we can't compute the required antiderivative. But looking at the series, it would appear that it must converge, because the terms we are adding are smaller than the terms of a p-series, that is,

$$\frac{1}{n^2 \ln n} < \frac{1}{n^2},$$

when $n \geq 3$. Since adding up the terms $1/n^2$ doesn't get "too big", the new series "should" also converge. Let's make this more precise.

The series $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n^2 \ln n}$ converges if and only if $\displaystyle\sum_{n=3}^{\infty} \frac{1}{n^2 \ln n}$ converges—all we've done is dropped the initial term. We know that $\displaystyle\sum_{n=3}^{\infty} \frac{1}{n^2}$ converges. Looking at two typical partial sums:

$$s_n = \frac{1}{3^2 \ln 3} + \frac{1}{4^2 \ln 4} + \frac{1}{5^2 \ln 5} + \cdots + \frac{1}{n^2 \ln n} < \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \cdots + \frac{1}{n^2} = t_n.$$

Since the p-series converges, say to L, and since the terms are positive, $t_n < L$. Since the terms of the new series are positive, the s_n form an increasing sequence and $s_n < t_n < L$ for all n. Hence the sequence $\{s_n\}$ is bounded and so converges. ♣

Like the Integral Test, the so-called **Comparison Test** can be used to show both convergence and divergence. In the case of the Integral Test, a single calculation will confirm whichever is the case. To use the Comparison Test we must first have a good idea as to convergence or divergence and pick the sequence for comparison accordingly.

Example 6.50: Divergence by Comparison

Does $\displaystyle\sum_{n=2}^{\infty} \frac{1}{\sqrt{n^2-3}}$ *converge?*

Solution. We observe that the -3 should have little effect compared to the n^2 inside the square root. Therefore, we guess that the terms are enough like $1/\sqrt{n^2} = 1/n$ so that the series should diverge. This analysis leads us to apply the Comparison Test based on the harmonic series. We note that

$$\frac{1}{\sqrt{n^2-3}} > \frac{1}{\sqrt{n^2}} = \frac{1}{n},$$

so that

$$s_n = \frac{1}{\sqrt{2^2-3}} + \frac{1}{\sqrt{3^2-3}} + \cdots + \frac{1}{\sqrt{n^2-3}} > \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} = t_n,$$

where t_n is 1 less than the corresponding partial sum of the harmonic series because we start at $n = 2$ instead of $n = 1$. Since $\lim\limits_{n\to\infty} t_n = \infty$, $\lim\limits_{n\to\infty} s_n = \infty$ as well. Hence, the given series diverges. ♣

For reference, we summarize the Comparison Test in a theorem.

Theorem 6.51: Comparison Test for Series

Suppose that a_n and b_n are non-negative for all n and that $a_n \leq b_n$ when $n \geq N$, for some N.

1. *If* $\displaystyle\sum_{n=0}^{\infty} b_n$ **converges**, *then* $\displaystyle\sum_{n=0}^{\infty} a_n$ *also* **converges**.
2. *If* $\displaystyle\sum_{n=0}^{\infty} a_n$ **diverges**, *then* $\displaystyle\sum_{n=0}^{\infty} b_n$ *also* **converges**.

Note:

1. Sometimes, even when the Integral Test applies, comparison to a known series is easier, so it's generally a good idea to think about doing a comparison before doing the Integral Test.
2. The general approach is this: If you believe that a new series is convergent, attempt to find a convergent series whose terms are larger than the terms of the new series; if you believe that a new series is divergent, attempt to find a divergent series whose terms are smaller than the terms of the new series.

Example 6.52: Applying the Comparison Test

Does $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n^2+3}}$ *converge?*

Solution. Just as in the last example, we guess that this is very much like the harmonic series and so diverges. Unfortunately,

$$\frac{1}{\sqrt{n^2+3}} < \frac{1}{n},$$

so we can't compare the series directly to the harmonic series. A little thought leads us to

$$\frac{1}{\sqrt{n^2+3}} > \frac{1}{\sqrt{n^2+3n^2}} = \frac{1}{2n},$$

so if $\sum 1/(2n)$ diverges then the given series diverges. But since $\sum 1/(2n) = (1/2)\sum 1/n$, Theorem 6.29 implies that it does indeed diverge. ♣

Example 6.53: Integral Test or Comparison Test

Does $\sum_{n=2}^{\infty} \frac{|\sin n|}{n^2}$ *converge?*

Solution. We can't apply the Integral Test here, because the terms of this series are not decreasing. Just as in Example 6.49, however,

$$\frac{|\sin n|}{n^2} \le \frac{1}{n^2},$$

because $|\sin n| \le 1$. Once again, the partial sums

$$s_n = \sum_{i=2}^{n} \frac{|\sin i|}{i^2}$$

are non-decreasing and bounded above by

$$t_n = \sum_{i=2}^{n} \frac{1}{i^2} < L,$$

where L is the sum of the p-series $\sum_{n=2}^{\infty} \frac{1}{n^2}$. So the given series converges. ♣

Exercises for Section 6.5

Exercise 6.5.1 *Determine whether the series converge or diverge.*

(a) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{2n^2+3n+5}$

(b) $\displaystyle\sum_{n=2}^{\infty} \frac{1}{2n^2+3n-5}$

(c) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{2n^2-3n-5}$

(d) $\displaystyle\sum_{n=1}^{\infty} \frac{3n+4}{2n^2+3n+5}$

(e) $\displaystyle\sum_{n=1}^{\infty} \frac{3n^2+4}{2n^2+3n+5}$

(f) $\displaystyle\sum_{n=1}^{\infty} \frac{\ln n}{n}$

(g) $\displaystyle\sum_{n=1}^{\infty} \frac{\ln n}{n^3}$

(h) $\displaystyle\sum_{n=2}^{\infty} \frac{1}{\ln n}$

(i) $\displaystyle\sum_{n=1}^{\infty} \frac{3^n}{2^n+5^n}$

(j) $\displaystyle\sum_{n=1}^{\infty} \frac{3^n}{2^n+3^n}$

6.6 Absolute and Conditional Convergence

Roughly speaking there are two ways for a series to converge: As in the case of $\sum 1/n^2$, the individual terms get small very quickly, so that the sum of all of them stays finite, or, as in the case of $\sum(-1)^{n-1}/n$, the terms don't get small fast enough ($\sum 1/n$ diverges), but a mixture of positive and negative terms provides enough cancellation to keep the sum finite. You might guess from what we've seen that if the terms get small fast enough that the sum of their absolute values converges, then the series will still converge regardless of which terms are actually positive or negative. This leads us to the following theorem.

Theorem 6.54: Absolute Convergence Test

If $\displaystyle\sum_{n=0}^{\infty} |a_n|$ converges, then $\displaystyle\sum_{n=0}^{\infty} a_n$ converges.

Proof. Note that $0 \le a_n + |a_n| \le 2|a_n|$ so by the Comparison Test $\displaystyle\sum_{n=0}^{\infty} (a_n + |a_n|)$ converges. Since

$$\sum_{n=0}^{\infty} (a_n + |a_n|) - \sum_{n=0}^{\infty} |a_n| = \sum_{n=0}^{\infty} a_n + |a_n| - |a_n| = \sum_{n=0}^{\infty} a_n,$$

by Theorem 6.29, the conclusion follows. ♣

Note:

1. Intuitively, this theorem says that it is (potentially) easier for $\sum a_n$ to converge than for $\sum |a_n|$ to converge, because terms may partially cancel in the first series.

2. So given a series $\sum a_n$ with both positive and negative terms, you should first ask whether $\sum |a_n|$ converges. This may be an easier question to answer, because we have tests that apply specifically to series with non-negative terms.

   - If $\sum |a_n|$ converges then you know that $\sum a_n$ converges as well.
   - If $\sum |a_n|$ diverges then it still may be true that $\sum a_n$ converges, but you will need to use other techniques to decide.

If $\sum |a_n|$ converges we say that $\sum a_n$ **converges absolutely**. To say that $\sum a_n$ converges absolutely is to say that the terms of the series get small (in absolute value) quickly enough to guarantee that the series converges, regardless of whether any of the terms cancel each other. For example $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n^2}$ converges absolutely.

Definition 6.55: Absolutely Convergent

Given a series $\sum_{n=1}^{\infty} a_n$. *If the corresponding series* $\sum_{n=1}^{\infty} |a_n|$ *converges, then* $\sum_{n=1}^{\infty} a_n$ **converges absolutely**.

If $\sum a_n$ converges but $\sum |a_n|$ does not, we say that $\sum a_n$ **converges conditionally**. Recall that the alternating harmonic series $\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}$ converges, but that the corresponding series of absolute values, namely the harmonic series $\sum_{n=1}^{\infty}\frac{1}{n}$, diverges. Hence, the alternating harmonic series is conditionally convergent.

Definition 6.56: Conditionally Convergent

Given a series $\sum_{n=1}^{\infty} a_n$. *If* $\sum_{n=1}^{\infty} a_n$ *converges, but the corresponding series* $\sum_{n=1}^{\infty} |a_n|$ *does not converge, then* $\sum_{n=1}^{\infty} a_n$ **converges conditionally**.

Note: Instead of writing that a series *converges absolutely* (or *conditionally*), we may also use the expression the series is **absolutely** (or **conditionally**) **convergent**.

Example 6.57: Absolutely Convergent, Conditionally Convergent, or Divergent

Does $\sum_{n=2}^{\infty}\frac{\sin n}{n^2}$ *converge? If so, how?*

Solution. In Example 6.53 we saw that $\sum_{n=2}^{\infty}\frac{|\sin n|}{n^2}$ converges, so the given series converges absolutely. ♣

Example 6.58: Absolutely Convergent, Conditionally Convergent, or Divergent

Does $\sum_{n=0}^{\infty}(-1)^n \frac{3n+4}{2n^2+3n+5}$ *converge? If so, how?*

Solution. We begin by analyzing the corresponding series of absolute values, where

$$a_n = \frac{3n+4}{2n^2+3n+5}.$$

Looking at the leading terms in the numerator and denominator of a_n, we speculate to compare this series with the harmonic series:

Left side	Right side
$\frac{1}{n}$	$\frac{3n+4}{2n^2+3n+5}$
$2n^2+3n+5$	$n(3n+4)$
$2n^2+3n+5$	$3n^2+4n$
5	n^2+n

Indeed, $n^2+n > 5$ for $n > 1$, and therefore this series diverges since the harmonic series diverges. So if the original series converges, it does so conditionally.

To test for convergence, we apply the Alternating Series Test:

1. Clearly, $a_n = \frac{3n+4}{2n^2+3n+5}$ is positive for all $n \geq 0$.

2. $\lim_{n\to\infty} \frac{3n+4}{2n^2+3n+5} = 0.$

3. If we let $f(x) = \frac{3x+4}{2x^2+3x+5}$ then $f'(x) = -\frac{6x^2+16x-3}{(2x^2+3x+5)^2}$, and it is not hard to see that this is negative for $x \geq 1$, so the sequence $\left\{\frac{3n+4}{2n^2+3n+5}\right\}_{n=0}^{\infty}$ is decreasing.

Thus, the series $\sum_{n=0}^{\infty}(-1)^n \frac{3n+4}{2n^2+3n+5}$ converges by the Alternating Series Test, and we conclude that the series converges conditionally. ♣

Exercises for Section 6.6

Exercise 6.6.1 *Determine whether each series converges absolutely, converges conditionally, or diverges.*

(a) $\displaystyle\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{2n^2+3n+5}$

(b) $\displaystyle\sum_{n=1}^{\infty}(-1)^{n-1}\frac{3n^2+4}{2n^2+3n+5}$

(c) $\displaystyle\sum_{n=1}^{\infty}(-1)^{n-1}\frac{\ln n}{n}$

(d) $\displaystyle\sum_{n=1}^{\infty}(-1)^{n-1}\frac{\ln n}{n^3}$

(e) $\displaystyle\sum_{n=2}^{\infty}(-1)^{n}\frac{1}{\ln n}$

(f) $\displaystyle\sum_{n=0}^{\infty}(-1)^{n}\frac{3^n}{2^n+5^n}$

(g) $\displaystyle\sum_{n=0}^{\infty}(-1)^{n}\frac{3^n}{2^n+3^n}$

(h) $\displaystyle\sum_{n=1}^{\infty}(-1)^{n-1}\frac{\arctan n}{n}$

6.7 Ratio and Root Tests

Does the series $\displaystyle\sum_{n=0}^{\infty}\frac{n^5}{5^n}$ converge? It is possible, but a bit unpleasant, to approach this with the Integral Test or the Comparison Test, but there is an easier way. Consider what happens as we move from one term to the next in this series:

$$\cdots+\frac{n^5}{5^n}+\frac{(n+1)^5}{5^{n+1}}+\cdots$$

The denominator goes up by a factor of 5, $5^{n+1}=5\cdot 5^n$, but the numerator goes up by much less: $(n+1)^5=n^5+5n^4+10n^3+10n^2+5n+1$, which is much less than $5n^5$ when n is large, because $5n^4$ is much less than n^5. So we might guess that in the long run it begins to look as if each term is $1/5$ of the previous term. We have seen series that behave like this: The geometric series.

$$\sum_{n=0}^{\infty}\frac{1}{5^n}=\frac{5}{4},$$

So we might try comparing the given series to some variation of this geometric series. This is possible, but a bit messy. We can in effect do the same thing, but bypass most of the unpleasant work.

The key is to notice that

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=\lim_{n\to\infty}\frac{(n+1)^5}{5^{n+1}}\frac{5^n}{n^5}=\lim_{n\to\infty}\frac{(n+1)^5}{n^5}\frac{1}{5}=1\cdot\frac{1}{5}=\frac{1}{5}.$$

This is really just what we noticed above, done a bit more formally: in the long run, each term is one fifth of the previous term. Now pick some number between $1/5$ and 1, say $1/2$. Because

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=\frac{1}{5},$$

then when n is big enough, say $n\ge N$ for some N,

$$\frac{a_{n+1}}{a_n} < \frac{1}{2} \quad \text{so} \quad a_{n+1} < \frac{a_n}{2}.$$

So $a_{N+1} < a_N/2$, $a_{N+2} < a_{N+1}/2 < a_N/4$, $a_{N+3} < a_{N+2}/2 < a_N/8$, and so on. The general form is $a_{N+k} < a_N/2^k$. So if we look at the series

$$\sum_{k=0}^{\infty} a_{N+k} = a_N + a_{N+1} + a_{N+2} + a_{N+3} + \cdots + a_{N+k} + \cdots,$$

its terms are less than or equal to the terms of the sequence

$$a_N + \frac{a_N}{2} + \frac{a_N}{4} + \frac{a_N}{8} + \cdots + \frac{a_N}{2^k} + \cdots = \sum_{k=0}^{\infty} \frac{a_N}{2^k} = 2a_N.$$

So by the Comparison Test, $\sum_{k=0}^{\infty} a_{N+k}$ converges, and this means that $\sum_{n=0}^{\infty} a_n$ converges, since we've just added the fixed number $a_0 + a_1 + \cdots + a_{N-1}$.

Under what circumstances could we do this? The crucial part was that the limit of a_{n+1}/a_n, say L, was less than 1 so that we could pick a value r so that $L < r < 1$. The fact that $L < r$ (in our example $1/5 < 1/2$) means that we can compare the series $\sum a_n$ to $\sum r^n$, and the fact that $r < 1$ guarantees that $\sum r^n$ converges. That's really all that is required to make the argument work.

Theorem 6.59: Ratio Test

Given a series $\sum a_n$ with positive terms and $\lim\limits_{n\to\infty} \frac{a_{n+1}}{a_n} = L$:

1. *If $L < 1$, then the series converges.*
2. *If $L > 1$, then the series diverges.*
3. *If $L = 1$, then this test gives no information.*

Proof. The example above essentially proves the first part of this, if we simply replace $1/5$ by L and $1/2$ by r. Suppose that $L > 1$, and pick r so that $1 < r < L$. Then for $n \geq N$, for some N,

$$\frac{|a_{n+1}|}{|a_n|} > r \quad \text{and} \quad |a_{n+1}| > r|a_n|.$$

This implies that $|a_{N+k}| > r^k|a_N|$, but since $r > 1$ this means that $\lim\limits_{k\to\infty} |a_{N+k}| \neq 0$, which means also that $\lim\limits_{n\to\infty} a_n \neq 0$. By the Divergence Test, the series diverges.

To see that we get no information when $L = 1$, we need to exhibit two series with $L = 1$, one that converges and one that diverges. The series $\sum 1/n^2$ and $\sum 1/n$ provide a simple example. ♣

Note:

1. The Ratio Test is particularly useful for series involving factorials and exponentials.

2. Absolute Convergence:

   - In general, we require a series to just have non-zero terms.
   - Then we consider the absolute values of the terms: $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = L$
   - This means we are testing for absolute convergence.

Example 6.60: Factorials and Ratio Test

Analyze $\sum_{n=0}^{\infty}\frac{5^n}{n!}$.

Solution.

$$\lim_{n\to\infty}\frac{5^{n+1}}{(n+1)!}\frac{n!}{5^n} = \lim_{n\to\infty}\frac{5^{n+1}}{5^n}\frac{n!}{(n+1)!} = \lim_{n\to\infty}5\frac{1}{(n+1)} = 0.$$

Since $0 < 1$, the series converges. ♣

A similar argument to the one used for the Ratio Test justifies a related test that is occasionally easier to apply, namely the so-called **Root Test**.

Theorem 6.61: Root Test

Given a series $\sum a_n$ *with positive terms and* $\lim_{n\to\infty}(a_n)^{1/n} = L$:

1. *If* $L < 1$, *then the series converges.*
2. *If* $L > 1$, *then the series diverges.*
3. *If* $L = 1$, *then this test gives no information.*

The proof of the Root Test is actually easier than that of the Ratio Test, and is left as an exercise.

Example 6.62: Exponentials and Root Test

Analyze $\sum_{n=0}^{\infty}\frac{5^n}{n^n}$.

Solution. The Ratio Test turns out to be a bit difficult on this series (try it). Using the Root Test:

$$\lim_{n\to\infty}\left(\frac{5^n}{n^n}\right)^{1/n} = \lim_{n\to\infty}\frac{(5^n)^{1/n}}{(n^n)^{1/n}} = \lim_{n\to\infty}\frac{5}{n} = 0.$$

Since $0 < 1$, the series converges. ♣

Note:

1. The Root Test is frequently useful for series involving exponentials.

2. Absolute Convergence:

   - In general, we require a series to just have non-zero terms.
   - Then we consider the absolute values of the terms: $\lim_{n\to\infty} (|a_n|)^{1/n} = L$
   - This means we are testing for absolute convergence.

Exercises for Section 6.7

Exercise 6.7.1 *Compute the following limit for the given series.*

(a) $\lim_{n\to\infty} |a_{n+1}/a_n|$ *for* $\sum 1/n^2$

(b) $\lim_{n\to\infty} |a_{n+1}/a_n|$ *for* $\sum 1/n$

(c) $\lim_{n\to\infty} |a_n|^{1/n}$ *for* $\sum 1/n^2$

(d) $\lim_{n\to\infty} |a_n|^{1/n}$ *for* $\sum 1/n$

Exercise 6.7.2 *Determine whether the series converge.*

(a) $\displaystyle\sum_{n=0}^{\infty} (-1)^n \frac{3^n}{5^n}$

(b) $\displaystyle\sum_{n=1}^{\infty} \frac{n!}{n^n}$

(c) $\displaystyle\sum_{n=1}^{\infty} \frac{n^5}{n^n}$

(d) $\displaystyle\sum_{n=1}^{\infty} \frac{(n!)^2}{n^n}$

6.8 Power Series and Polynomial Approximation

In this chapter we have a closer look at so-called **power series**, which arise in the study of analytic functions. A power series is basically an *infinite degree polynomial* that represents some function. Since we know a lot more about polynomial functions than arbitrary functions, this allows us to readily differentiate, integrate, and approximate some functions using power series.

6.8.1. Power Series

Recall that the sum of a geometric series can be expressed using the simple formula:

$$\sum_{n=0}^{\infty} kx^n = \frac{k}{1-x},$$

if $|x| < 1$, and that the series diverges when $|x| \geq 1$. At the time, we thought of x as an unspecified constant, but we could just as well think of it as a variable, in which case the series

$$\sum_{n=0}^{\infty} kx^n$$

is a function, namely, the function $k/(1-x)$, as long as $|x| < 1$: Looking at this from the opposite perspective, this means that the function $k/(1-x)$ can be represented as the sum of an infinite series. Why would this be useful? While $k/(1-x)$ is a reasonably easy function to deal with, the more complicated representation $\sum kx^n$ does have some advantages: it appears to be an infinite version of one of the simplest function types – a polynomial. Later on we will investigate some of the ways we can take advantage of this 'infinite polynomial' representation, but first we should ask if other functions can even be represented this way.

The geometric series has a special feature that makes it unlike a typical polynomial—the coefficients of the powers of x are all the same, namely k. We will need to allow more general coefficients if we are to get anything other than the geometric series.

Definition 6.63: Power Series Centred Around Zero

A **power series** *is a series of the form*

$$P(x) = \sum_{n=0}^{\infty} a_n x^n,$$

where the **coefficients** a_n *are real numbers.*

Note:

1. As we did in the section on sequences, we can think of the a_n as being a function $a(n)$ defined on the non-negative integers.
2. It is important to remember that the a_n do not depend on x.

Example 6.64: Power Series Convergence

Determine the values of x for which the power series $\displaystyle\sum_{n=1}^{\infty} \frac{x^n}{n}$ *converges.*

Solution. We can investigate convergence using the Ratio Test:

$$\lim_{n\to\infty} \frac{|x|^{n+1}}{n+1} \frac{n}{|x|^n} = \lim_{n\to\infty} |x| \frac{n}{n+1} = |x|.$$

Thus when $|x| < 1$ the series converges and when $|x| > 1$ it diverges, leaving only two values in doubt. When $x = 1$ the series is the harmonic series and diverges; when $x = -1$ it is the alternating harmonic series (actually the negative of the usual alternating harmonic series) and converges. Thus, we may think of $\sum_{n=1}^{\infty} \frac{x^n}{n}$ as a function from the interval $[-1,1)$ to the real numbers. ♣

A bit of thought reveals that the Ratio Test applied to a power series will always have the same nice form. In general, we will compute

$$\lim_{n\to\infty} \frac{|a_{n+1}||x|^{n+1}}{|a_n||x|^n} = \lim_{n\to\infty} |x| \frac{|a_{n+1}|}{|a_n|} = |x| \lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = L|x|,$$

assuming that $\lim |a_{n+1}|/|a_n|$ exists:

1. If $L \in (0,\infty)$:
   - Then the series converges if $L|x| < 1$, that is, if $|x| < 1/L$, and diverges if $|x| > 1/L$.
   - Only the two values $x = \pm 1/L$ require further investigation.
   - The value $1/L$ is called the **radius of convergence**.
   - Thus the series will always define a function on the interval $(-1/L, 1/L)$, that perhaps will extend to one or both endpoints as well.
   - This interval is referred to as the **interval of convergence**.
   - This interval is essentially the domain of the power series.
2. If $L = 0$:
   - Then no matter what value x takes the limit is 0.
   - The series converges for all x and the function is defined for all real numbers.
3. If $L = \infty$:
   - Then no matter what value x takes the limit is infinite.
   - The series converges only when $x = 0$.

We can make these ideas a bit more general. Consider the series

$$\sum_{n=0}^{\infty} \frac{(x+2)^n}{3^n}$$

This looks a lot like a power series, but with $(x+2)^n$ instead of x^n. Let's try to determine the values of x for which it converges. This is just a geometric series, so it converges when

$$\begin{aligned} |x+2|/3 &< 1 \\ |x+2| &< 3 \\ -3 < x+2 &< 3 \\ -5 < x &< 1. \end{aligned}$$

So the interval of convergence for this series is $(-5,1)$. The centre of this interval is at -2, which is at distance 3 from the endpoints, so the radius of convergence is 3, and we say that the series is centred at -2.

Interestingly, if we compute the sum of the series we get

$$\sum_{n=0}^{\infty}\left(\frac{x+2}{3}\right)^n = \frac{1}{1-\frac{x+2}{3}} = \frac{3}{1-x}.$$

Multiplying both sides by 1/3 we obtain

$$\sum_{n=0}^{\infty}\frac{(x+2)^n}{3^{n+1}} = \frac{1}{1-x},$$

which we recognize as being equal to

$$\sum_{n=0}^{\infty} x^n,$$

so we have two series with the same sum but different intervals of convergence.

This leads to the following definition:

Definition 6.65: Power Series Centred Around a

*A **power series** centred at a has the form*

$$P(x) = \sum_{n=0}^{\infty} a_n(x-a)^n,$$

*where the **centre** a and **coefficients** a_n are real numbers.*

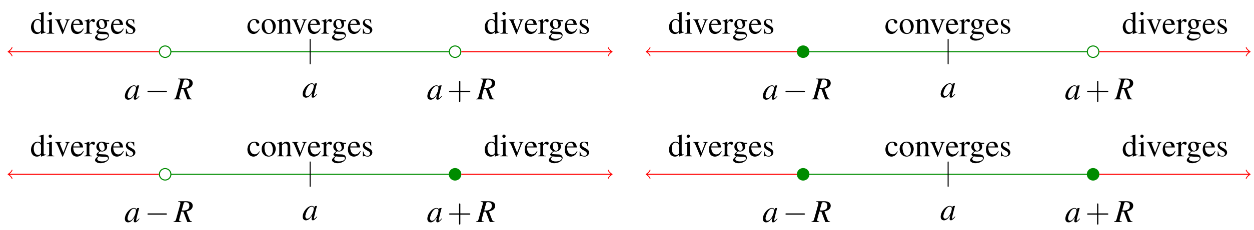
Note: The power series centred at zero given in Definition 6.63 is a special case of the above definition when $a = 0$.

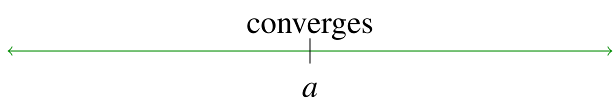
Convergence of a Power Series

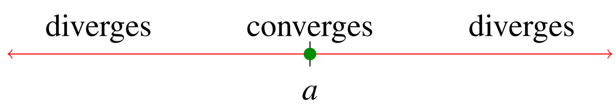
Given a power series $\sum a_n(x-a)^n$ and its radius of convergence R, the series behaves in one of three ways:

1. The series converges absolutely for x with $|x-a| < R$, it diverges for x with $|x-a| > R$, and at $x = a-R$ and $x = a+R$ further investigation is needed.
2. When $R = \infty$, the series converges absolutely for every x.
3. When $R = 0$, the series converges at $x = a$ and diverges everywhere else.

The convergence and divergence can be visualized:

1. 

2. 

3. 

Example 6.66: Interval of Convergence

Given the power series

$$\sum_{n=0}^{\infty} \frac{(-1)^n n(x-2)^n}{3^n}$$

determine the following:

(a) *radius of convergence* (b) *interval of convergence*

Solution. Obviously, the series converges for $x = 2$. To determine all values of x for which the series converges, we begin by applying the Ratio Test:

$$\begin{aligned}
L &= \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| \\
&= \lim_{n\to\infty} \left| \frac{(-1)^{n+1}(n+1)(x-2)^{n+1}}{3^{n+1}} \frac{3^n}{(-1)^n n(x-2)^n} \right| \\
&= \lim_{n\to\infty} \left| \frac{(n+1)(x-2)}{3n} \right| \\
&= |x-2| \lim_{n\to\infty} \frac{n+1}{3n} \\
&= \frac{1}{3}|x-2|
\end{aligned}$$

(a) By the Ratio Test, the radius of convergence is $R = 3$.

(b) We now determine the interval of convergence. By the Ratio Test, the series converges absolutely if $L < 1$:

$$\begin{aligned}
&\frac{1}{3}|x-2| < 1 \\
&\implies |x-2| < 3 \\
&\implies -3 < x-2 < 3 \\
&\implies -1 < x < 5
\end{aligned}$$

The series diverges if $L > 1$, i.e. $x < -1$ and $x > 5$.

Let us now look at the case when $L = 1$, which means investigating the behaviour of the series at endpoints $x = -1$ and $x = 5$:

Case $x = -1$: Then the series becomes

$$\sum_{n=0}^{\infty} \frac{(-1)^n n(-1-2)^n}{3^n} = \sum_{n=0}^{\infty} (-1)^{2n} n = \sum_{n=0}^{\infty} n.$$

Since $\lim_{n\to\infty} n = \infty \neq 0$, this series is divergent by the n-th Term Test (Divergence Test).

Case $x = 5$: Then the series becomes

$$\sum_{n=0}^{\infty} \frac{(-1)^n n(5-2)^n}{3^n} = \sum_{n=0}^{\infty} (-1)^n n.$$

Since $\lim_{n\to\infty} (-1)^n n$ does not exist, this series is also divergent by the n-th Term Test (Divergence Test).

Thus, the interval of convergence for the given power series is $x \in (-1, 5)$.

♣

Example 6.67: Interval of Convergence

Given the power series

$$\sum_{n=1}^{\infty} (n+1)!(x+3)^n$$

determine the following:

(a) radius of convergence *(b) interval of convergence*

Solution. Obviously, the series converges for $x = -3$. To determine all values of x for which the series converges, we begin by applying the Ratio Test:

$$\begin{aligned} L &= \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| \\ &= \lim_{n\to\infty} \left| \frac{(n+2)!(x+3)^{n+1}}{(n+1)!(x+3)^n} \right| \\ &= |x+3| \lim_{n\to\infty} (n+1) = \infty, \end{aligned}$$

provided that $x \neq -3$. Therefore, this series will only converge for $x = 3$.

(a) The radius of convergence is $R = 0$.

(b) The interval of convergence is $x = -3$, which is just one point.

♣

Exercises for Section 6.8.1

Exercise 6.8.1 *Find the radius and interval of convergence for each series. In part c), do not attempt to determine whether the endpoints are in the interval of convergence.*

(a) $\displaystyle\sum_{n=0}^{\infty} nx^n$

(b) $\displaystyle\sum_{n=0}^{\infty} \frac{x^n}{n!}$

(c) $\displaystyle\sum_{n=1}^{\infty} \frac{n!}{n^n}(x-2)^n$

(d) $\displaystyle\sum_{n=1}^{\infty} \frac{(n!)^2}{n^n}(x-2)^n$

(e) $\displaystyle\sum_{n=1}^{\infty} \frac{(x+5)^n}{n(n+1)}$

(f) $\displaystyle\sum_{n=1}^{\infty} \frac{n!}{n^n}x^n$

6.8.2. Calculus with Power Series

We now know that some functions can be expressed as power series, which look like infinite polynomials. In fact, there are many extremely useful functions that cannot be represented using the usual elementary functions, but they can be represented using a power series. For example, the Bessel functions and hypergeometric functions which are important in such fields as engineering, physics and acoustics. Since it is easy to find derivatives and integrals of polynomials, we might hope that we can take derivatives and integrals of power series in an analogous way. In fact we can, as stated in the following theorem, which we will not prove here.

Theorem 6.68: Differentiation and Integration of a Power Series

Suppose the power series $f(x) = \displaystyle\sum_{n=0}^{\infty} a_n(x-a)^n$ *has radius of convergence* R.

1. *The* **derivative** *of* f *is the term by term differentiation of the power series:*

$$f'(x) = \sum_{n=1}^{\infty} na_n(x-a)^{n-1}$$

   with radius of convergence R.

2. *The* **integral** *of* f *is the term by term integration of the power series:*

$$\int f(x)\,dx = C + \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-a)^{n+1}$$

   with radius of convergence R.

Note: The process of differentiation and integration allows us to create further power series representations of functions.

Example 6.69: Power Series Representation

Given $f(x) = \ln|1-x|$*:*

(a) Find a power series representation of f*.*

(b) Use the first seven terms of this series to approximate $\ln(3/2)$*.*

(c) Approximate $\ln(9/4)$*.*

Solution.

(a) We recall that the derivative of $\ln|1-x|$ is $\frac{1}{1-x}$, and so we start with the geometric series:

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$$

when $|x| < 1$. Next, we integrate to obtain the function f:

$$\int \frac{1}{1-x}\,dx = -\ln|1-x| = \sum_{n=0}^{\infty} \frac{1}{n+1} x^{n+1}$$

$$f(x) = \ln|1-x| = \sum_{n=0}^{\infty} -\frac{1}{n+1} x^{n+1},$$

when $|x| < 1$. The series does not converge when $x = 1$, since $\sum_{n=0}^{\infty} -\frac{1}{n+1}(1)^{n+1} = -\sum_{n=1}^{\infty} \frac{1}{n}$ is the negative harmonic series. But the series does converge when $x = -1$, since $\sum_{n=0}^{\infty} -\frac{1}{n+1}(-1)^{n+1} = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ is the alternating harmonic series. Therefore, the interval of convergence is $[-1, 1)$.

(b) We note that $-1 \le x < 1$ implies $0 < 1-x \le 2$, and that $0 < \frac{3}{2} \le 2$, so we need $x = -\frac{1}{2}$. Then

$$\ln(3/2) = \ln|-(-1/2)| = \sum_{n=0}^{\infty} -\frac{1}{n+1}\left(-\frac{1}{2}\right)^{n+1} = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(n+1)2^{n+1}}.$$

Now we use the first seven terms to approximate $\ln(3/2)$:

$$\ln(3/2) \approx \frac{1}{2} - \frac{1}{8} + \frac{1}{24} - \frac{1}{64} + \frac{1}{160} - \frac{1}{384} + \frac{1}{896} = \frac{909}{2240} \approx 0.406.$$

Because this is an alternating series with decreasing terms, we know that the true value is between $909/2240$ and $909/2240 - 1/2048 = 29053/71680 \approx .4053$, so $0.4053 \le \ln(3/2) \le 0.406$.

(c) With a bit of arithmetic, we can approximate values outside of the interval of convergence, such as $9/4 > 2$. We can use the approximation we just computed, plus some rules for logarithms:

$$\ln(9/4) = \ln((3/2)^2) = 2\ln(3/2) \approx 0.812,$$

and using our bounds above,

$$0.8106 \le \ln(9/4) \le 0.812.$$

♣

Exercises for Section 6.8.2

Exercise 6.8.2 *Find a series representation for*

(a) $\ln 2$

(b) $\ln(9/2)$

Exercise 6.8.3 *Find a power series representation for the following functions.*

(a) $1/(1-x)^2$

(b) $2/(1-x)^3$

(c) $1/(1-x)^3$

(d) $\displaystyle\int \ln(1-x)\,dx$

6.8.3. Maclaurin Series and Taylor Series

We have seen that some functions can be represented as series, which may give valuable information about the function. So far, we have seen only those examples that result from manipulation of our one fundamental example, the geometric series. We would like to start with a given function and produce a series to represent it, if possible.

Suppose that $f(x) = \sum_{n=0}^{\infty} a_n x^n$ on some interval of convergence centred at 0. Then we know that we can compute derivatives of f by taking derivatives of the terms of the series. Let's compute the first few derivatives and look for a pattern:

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} = a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + \cdots$$

$$f''(x) = \sum_{n=2}^{\infty} n(n-1) a_n x^{n-2} = 2a_2 + 3\cdot 2a_3 x + 4\cdot 3a_4 x^2 + \cdots$$

$$f'''(x) = \sum_{n=3}^{\infty} n(n-1)(n-2) a_n x^{n-3} = 3\cdot 2a_3 + 4\cdot 3\cdot 2a_4 x + \cdots$$

By examining these it's not hard to discern the general pattern. The kth derivative must be

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)(n-2)\cdots(n-k+1) a_n x^{n-k}$$

$$= k(k-1)(k-2)\cdots(2)(1)a_k + (k+1)(k)\cdots(2)a_{k+1}x$$
$$+ (k+2)(k+1)\cdots(3)a_{k+2}x^2 + \cdots$$

We can express this more clearly by using factorial notation:

$$f^{(k)}(x) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} a_n x^{n-k} = k!a_k + (k+1)!a_{k+1}x + \frac{(k+2)!}{2!} a_{k+2}x^2 + \cdots$$

We can solve for a_k by substituting $x = 0$ in the formula for $f^{(k)}(x)$:

$$f^{(k)}(0) = k!a_k + \sum_{n=k+1}^{\infty} \frac{n!}{(n-k)!} a_n 0^{n-k} = k!a_k,$$

$$a_k = \frac{f^{(k)}(0)}{k!}.$$

Note that the original series for f yields $f(0) = a_0$. So if a function f can be represented by a series, we can easily find such a series. This leads us to the definition of the so-called **Maclaurin series**.

Definition 6.70: Maclaurin Series

Suppose a function f is defined at $x = 0$ and whose derivatives all exist at $x = 0$. Then the **Maclaurin series** *of f at $x = 0$ is given by*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$
$$= f(0) + f'(0) + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \cdots$$

Note: A warning is in order here. Given a function f we may be able to compute the Maclaurin series, but that does not mean we have found a series representation for f. We still need to know where the series converges, and if, where it converges, it converges to $f(x)$. While for most commonly encountered functions the Maclaurin series does indeed converge to f on some interval, this is not true of all functions, so care is required.

Example 6.71: Maclaurin Series

Find the Maclaurin series for $f(x) = 1/(1-x)$.

Solution. We need to compute the derivatives of f and hope to spot a pattern.

$$f(x) = (1-x)^{-1}$$
$$f'(x) = (1-x)^{-2}$$
$$f''(x) = 2(1-x)^{-3}$$

$$
\begin{aligned}
f'''(x) &= 6(1-x)^{-4} \\
f^{(4)}(x) &= 4!(1-x)^{-5} \\
&\vdots \\
f^{(n)}(x) &= n!(1-x)^{-n-1}
\end{aligned}
$$

So

$$a_n = \frac{f^{(n)}(0)}{n!} = \frac{n!(1-0)^{-n-1}}{n!} = 1$$

and the Maclaurin series is

$$\sum_{n=0}^{\infty} 1 \cdot x^n = \sum_{n=0}^{\infty} x^n,$$

the geometric series with interval of convergence $|x| < 1$. ♣

As a practical matter, if we are interested in using a series to approximate a function, we will need some finite number of terms of the series. Even for functions with messy derivatives we can compute these using computer software like Sage. If we want to describe a series completely, we would like to be able to write down a formula for a typical term in the series. Fortunately, a few of the most important functions are very easy.

Example 6.72: Maclaurin Series

Find the Maclaurin series for $f(x) = \sin x$.

Solution. Computing the first few derivatives is simple: $f'(x) = \cos x$, $f''(x) = -\sin x$, $f'''(x) = -\cos x$, $f^{(4)}(x) = \sin x$, and then the pattern repeats. The values of the derivative when $x = 0$ are: 1, 0, -1, 0, 1, 0, -1, 0,..., and so the Maclaurin series is

$$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}.$$

We should always determine the radius of convergence:

$$\lim_{n\to\infty} \frac{|x|^{2n+3}}{(2n+3)!} \frac{(2n+1)!}{|x|^{2n+1}} = \lim_{n\to\infty} \frac{|x|^2}{(2n+3)(2n+2)} = 0,$$

so the series converges for every x. Since it turns out that this series does indeed converge to $\sin x$ everywhere, we have a series representation for $\sin x$ for every x. ♣

Sometimes the formula for the n-th derivative of a function f is difficult to discover, but a combination of a known Maclaurin series and some algebraic manipulation leads easily to the Maclaurin series for f.

Example 6.73: Maclaurin Series

Find the Maclaurin series for $f(x) = x\sin(-x)$.

Solution. To get from $\sin x$ to $x\sin(-x)$ we substitute $-x$ for x and then multiply by x. We can do the same thing to the series for $\sin x$:

$$x\sum_{n=0}^{\infty}(-1)^n\frac{(-x)^{2n+1}}{(2n+1)!} = x\sum_{n=0}^{\infty}(-1)^n(-1)^{2n+1}\frac{x^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty}(-1)^{n+1}\frac{x^{2n+2}}{(2n+1)!}.$$

♣

As we have seen, a power series can be centred at a point other than zero, and the method that produces the Maclaurin series can also produce such series, which are referred to as **Taylor series**.

Definition 6.74: Taylor Series

Suppose a function f is defined at $x=a$ and whose derivatives all exist at $x=a$. Then the **Taylor series** *of f centred at $x=a$ is given by*

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x-a)^n \\ &= f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\cdots+\frac{f^{(n)}(a)}{n!}(x-a)^n+\cdots \end{aligned}$$

Note: Notice that the Maclaurin series is a special case of the Taylor series when $a=0$.

Example 6.75: Taylor Series

Find a Taylor series centred at -2 for $f(x)=1/(1-x)$.

Solution.

Method 1: We develop the series from scratch. Suppose $f(x)=\sum_{n=0}^{\infty}a_n(x+2)^n$. We compute the k-th derivatives of f and the power series representation to solve for a_n:

Using our work in Example 6.71, the k-th derivative of $f(x)=\frac{1}{1-x}$ is given by

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

Based on our work at the beginning of this section, the k-th derivative of the power series representation is given by

$$\sum_{n=k}^{\infty}\frac{n}{(n-k)!}a_n(x+2)^{n-k}.$$

Now we equate these two results to obtain the following equation:

$$k!(1-x)^{-k-1}=\sum_{n=k}^{\infty}\frac{n!}{(n-k)!}a_n(x+2)^{n-k}$$

and substituting $x=-2$ we get $k!3^{-k-1}=k!a_k$ and $a_k=3^{-k-1}=1/3^{k+1}$, so the series is

$$\sum_{n=0}^{\infty}\frac{(x+2)^n}{3^{n+1}}.$$

Method 2: We use the formula for the Taylor series representation of f:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(-2)}{n!}(x+2)^n.$$

Now we compute the first few derivatives and evaluate them at $x = -2$ to deduce a pattern:

$$\begin{aligned} f(x) &= (1-x)^{-1} & f(-2) &= \frac{1}{3} \\ f'(x) &= (1-x)^{-2} & f'(-2) &= \frac{1}{3^2} \\ f''(x) &= 2!(1-x)^{-3} & f''(-2) &= \frac{3!}{3^4} \\ f'''(x) &= 3!(1-x)^{-4} & f'''(-2) &= \frac{3!}{3^4} \end{aligned}$$

Hence, the n-th derivative evaluated at $x = -2$ is

$$f^{(n)}(-2) = \frac{n!}{3^{n+1}},$$

and so, as before, the series is

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(-2)}{n!}(x+2)^n = \sum_{n=0}^{\infty} \frac{\frac{n!}{3^{n+1}}}{n!}(x+2)^n = \sum_{n=0}^{\infty} \frac{(x+2)^n}{3^{n+1}}.$$

♣

Example 6.76: Taylor Series

Given $f(x) = e^x$, find the following:

(a) The Taylor series of f centred at 3. *(b) The interval of convergence for the series.*

Solution.

(a) We use the formula for the Taylor series representation of f:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(3)}{n!}(x-3)^n$$

Since the derivative of $f(x) = e^x$ is again e^x, we have that the n-th derivative evaluated at $x = -3$ is

$$f^{(n)}(3) = e^3,$$

and so the series is

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(3)}{n}(x-3)^n = \sum_{n=0}^{\infty} \frac{e^3}{n!}(x-3)^n.$$

(b) We use the Ratio Test to determine the radius of convergence:

$$\begin{aligned} L &= \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| \\ &= \lim_{n\to\infty}\left|\frac{e^3(x-3)^{n+1}}{(n+1)!}\frac{n!}{e^3(x-3)^n}\right| \\ &= \lim_{n\to\infty}\left|\frac{x-3}{n+1}\right| \\ &= |x-3|\lim_{n\to\infty}\frac{1}{n+1} \\ &= 0 \end{aligned}$$

Hence, the radius of convergence is infinity, and so the series converges absolutely for every x, i.e. the interval of convergence is $(-\infty,\infty)$.

♣

Table 6.1 is a list of the five most common Maclaurin series and their interval of convergence.

Function	*Maclaurin series*	*Interval of convergence*
$\frac{1}{1-x}$	$\sum_{n=0}^{\infty} x^n$	$(-1,1)$
e^x	$\sum_{n=0}^{\infty} \frac{1}{n!}x^n$	$(-\infty,\infty)$
$\ln(1+x)$	$\sum_{n=1}^{\infty}(-1)^{n-1}\frac{x^n}{n}$	$(-1,1]$
$\sin(x)$	$\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!}$	$(-\infty,\infty)$
$\cos(x)$	$\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n)!}$	$(-\infty,\infty)$

Table 6.1: Common Maclaurin series.

Exercises for Section 6.8.3

Exercise 6.8.4 *For each function, find the Maclaurin series and the radius of convergence.*

(a) $f(x) = \cos x$

(b) $f(x) = e^x$

(c) $f(x) = \dfrac{1}{1+x}$

(d) $f(x) = \sin(2x)$

(e) $f(x) = \dfrac{1}{\sqrt{1-x}}$

Exercise 6.8.5 *For each function, find the Taylor series centred at a and the radius of convergence.*

(a) $f(x) = \dfrac{1}{x}$, $a = 5$

(b) $f(x) = \ln x$, $a = 1$

(c) $f(x) = \ln x$, $a = 2$

(d) $f(x) = \dfrac{1}{x^2}$, $a = 1$

Exercise 6.8.6 *Find the first four terms of the Maclaurin series for the following functions:*

(a) $f(x) = \tan x$

(b) $f(x) = x\cos(x^2)$

(c) $f(x) = xe^{-x}$

6.8.4. Taylor Polynomials

While the Taylor series can be thought of as an *infinite degree polynomial*, its partial sum s_n is a polynomial of degree n and referred to as **Taylor polynomial**.

Definition 6.77: Taylor Polynomial

Suppose a function f is defined at $x = a$. Then the **Taylor polynomial** *of f centred at $x = a$ with degree n is given by*

$$T_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Note: Notice that the first degree Taylor polynomial is the familiar linearization of f at $x = a$:

$$T_1(x) = f(a) + f'(a)(x-a) = L(x).$$

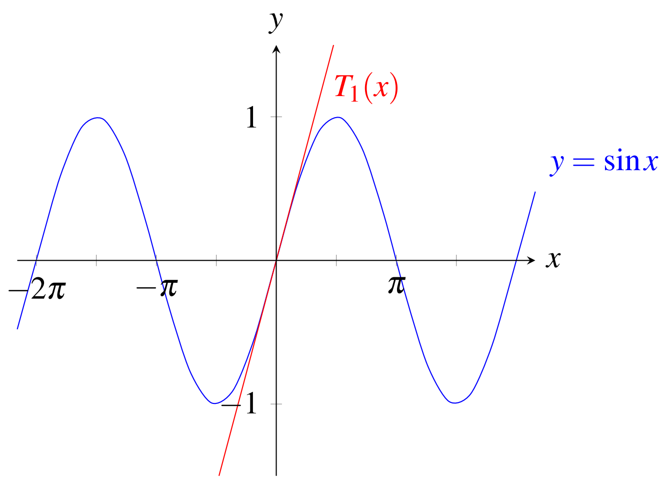
In fact, Taylor polynomials can be thought of as approximations of the function f. Let us explore this concept visually. In Example 6.72, we developed the Taylor series of $f(x) = sin(x)$ centred at zero

$$\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

with interval of convergence $(-\infty, \infty)$. A linear approximation of sine is given by

$$T_1(x) = f(0) + f'(0)(x-0) = cos(0)x = x$$

and shown below. Only at values close around zero is this linear approximation good enough to represent sine values.



When we calculate T_2,

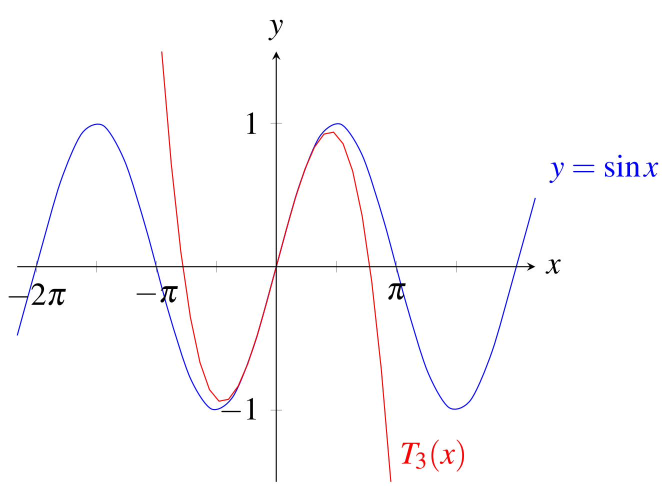
$$T_2(x) = T_1(x) + \frac{-\sin(0)}{2!} = T_1(x),$$

we notice that this is again T_1, since $-\sin(0) = 0$. In fact, $T_{2n} = T_{2n-1}$ for $n \geq 1$, since all even derivatives of the sine function will be positive or negative sine, and sine evaluated at zero is zero. This means that there is no quadratic Taylor polynomial approximation for sine using this series.

Let us compute T_3:

$$T_3(x) = T_1(x) + \frac{-\cos(0)}{3!}x^3 = x - \frac{x^3}{6},$$

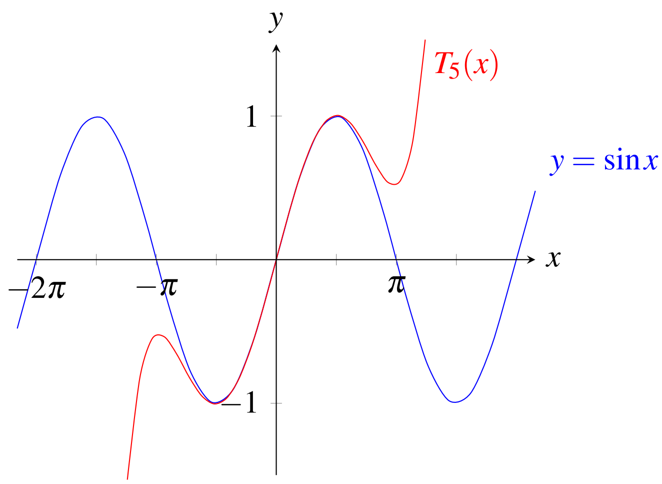
which yields a cubic approximation for sine as shown below. Notice that this cubic is already a better fit to sine than the linear one.



We compute one last Taylor polynomial, T_5:

$$T_5(x) = T_3(x) + \frac{\cos(0)}{5!}x^5 = x - \frac{x^3}{6} + \frac{x^5}{120}.$$

The graph below shows this approximation, and we can readily believe that continuing this process, we get an even better fit to sine, and thus a better and better approximation of sine.



Note: One question that naturally arises is, *How accurately do the Taylor polynomials approximate the function?* This question will be responded to in the next section.

Here is an example.

Example 6.78: Approximate e using Taylor Polynomials

Approximate e^x using Taylor polynomials at $a = 0$, and use this to approximate e.

Solution. In this case we use the function $f(x) = e^x$ at $a = 0$, and therefore

$$T_n(x) = a_0 + a_1 x + a_x x^2 + a_3 x^3 + \ldots + a_n x^n$$

Since all derivatives $f^{(k)}(x) = e^x$, we get:

$$\begin{aligned}
&a_0 = f(0) = 1 \\
&a_1 = \tfrac{f'(0)}{1!} = 1 \\
&a_2 = \tfrac{f''(0)}{2!} = \tfrac{1}{2!} \\
&a_3 = \tfrac{f'''(0)}{3!} = \tfrac{1}{3!} \\
&\ldots \\
&a_k = \tfrac{f^{(k)}(0)}{k!} = \tfrac{1}{k!} \\
&\ldots \\
&a_n = \tfrac{f^{(n)}(0)}{n!} = \tfrac{1}{n!}
\end{aligned}$$

Thus

$$\begin{aligned}
&T_1(x) = 1 + x = L(x) \\
&T_2(x) = 1 + x + \tfrac{x^2}{2!} \\
&T_3(x) = 1 + x + \tfrac{x^2}{2!} + \tfrac{x^3}{3!}
\end{aligned}$$

and in general

$$T_n(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}.$$

Finally, we can approximate $e = f(1)$ by simply calculating $T_n(1)$. A few values are:

$$\begin{aligned}
&T_1(1) = 1 + 1 = 2 \\
&T_2(1) = 1 + 1 + \tfrac{1^2}{2!} = 2.5 \\
&T_4(1) = 1 + 1 + \tfrac{1^2}{2!} + \tfrac{1^3}{3!} = 2.\overline{6} \\
&T_8(1) = 2.71825396825 \\
&T_{20}(1) = 2.71828182845
\end{aligned}$$

We can continue this way for larger values of n, but $T_{20}(1)$ is already a pretty good approximation of e, and we took only 20 terms! ♣

Exercises for Section 6.8.4

Exercise 6.8.7 *Find the 5^{th} degree Taylor polynomial for $f(x) = \sin x$ around $a = 0$.*

(a) Use this Taylor polynomial to approximate $\sin(0.1)$.

(b) Use a calculator to find $\sin(0.1)$. *How does this compare to our approximation in part (a)?*

Exercise 6.8.8 *Find the 3^{rd} degree Taylor polynomial for $f(x) = \frac{1}{1-x} - 1$ around $a = 0$. Explain why this approximation would not be useful for calculating $f(5)$.*

Exercise 6.8.9 *Consider $f(x) = \ln x$ around $a = 1$.*

(a) Find a general formula for $f^{(n)}(x)$ for $n \geq 1$.

(b) Find a general formula for the Taylor Polynomial, $T_n(x)$.

Exercise 6.8.10 *Approximate* $\ln(1.3)$ *to accuracy of at least 0.0001.*

6.8.5. Taylor's Theorem

One of the most important uses of infinite series is using an initial portion of the series for f to approximate f. We have seen, for example, that when we add up the first n terms of an alternating series with decreasing terms that the difference between this and the true value is at most the size of the next term. A similar result is true of many Taylor series, which is captured in **Taylor's Theorem**. This theorem is essentially a generalized version of the mean value theorem. It comes in a few different forms and we only state one form that is useful for this course.

Theorem 6.79: Taylor's Theorem

Suppose that f is defined on some open interval I around a and suppose $f^{(N+1)}(x)$ exists on this interval. Then for each $x \neq a$ in I there is a value z between x and a so that

$$f(x) = \sum_{n=0}^{N} \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(N+1)}(z)}{(N+1)!}(x-a)^{N+1}.$$

Proof. The proof requires some cleverness to set up, but then the details are quite elementary. We define a function $F(t)$ as follows:

$$F(t) = \sum_{n=0}^{N} \frac{f^{(n)}(t)}{n!}(x-t)^n + B(x-t)^{N+1}.$$

Here we have replaced a by t in the first $N+1$ terms of the Taylor series, and added a carefully chosen term on the end, with B to be determined. Note that we are temporarily keeping x fixed, so the only variable in this equation is t, and we will be interested only in t between a and x. Now substitute $t = a$:

$$F(a) = \sum_{n=0}^{N} \frac{f^{(n)}(a)}{n!}(x-a)^n + B(x-a)^{N+1}.$$

Set this equal to $f(x)$:

$$f(x) = \sum_{n=0}^{N} \frac{f^{(n)}(a)}{n!}(x-a)^n + B(x-a)^{N+1}.$$

Since $x \neq a$, we can solve this for B, which is a "constant"—it depends on x and a but those are temporarily fixed. Now we have defined a function $F(t)$ with the property that $F(a) = f(x)$. Also, all terms with a positive power of $(x-t)$ become zero when we substitute x for t, so $F(x) = f^{(0)}(x)/0! = f(x)$. So $F(a) = F(x)$. By Rolle's Theorem (from Differential Calculus), we know that there is a value $z \in (a,x)$ such that $F'(z) = 0$. But what is F'? Each term in $F(t)$, except the first term and the extra term involving B, is a product, so to take the derivative we use the Product Rule on each of these terms.

$$\begin{aligned} F(t) = f(t) &+ \frac{f^{(1)}(t)}{1!}(x-t)^1 + \frac{f^{(2)}(t)}{2!}(x-t)^2 + \frac{f^{(3)}(t)}{3!}(x-t)^3 + \cdots \\ &+ \frac{f^{(N)}(t)}{N!}(x-t)^N + B(x-t)^{N+1}. \end{aligned}$$

So the derivative is

$$\begin{aligned} F'(t) = f'(t) &+ \left(\frac{f^{(1)}(t)}{1!}(x-t)^0(-1) + \frac{f^{(2)}(t)}{1!}(x-t)^1\right) \\ &+ \left(\frac{f^{(2)}(t)}{1!}(x-t)^1(-1) + \frac{f^{(3)}(t)}{2!}(x-t)^2\right) \\ &+ \left(\frac{f^{(3)}(t)}{2!}(x-t)^2(-1) + \frac{f^{(4)}(t)}{3!}(x-t)^3\right) + \ldots \\ &+ \left(\frac{f^{(N)}(t)}{(N-1)!}(x-t)^{N-1}(-1) + \frac{f^{(N+1)}(t)}{N!}(x-t)^N\right) \end{aligned}$$

$$+B(N+1)(x-t)^N(-1).$$

The second term in each parenthesis cancel with the first term in the next one, leaving just

$$F'(t)=\frac{f^{(N+1)}(t)}{N!}(x-t)^N+B(N+1)(x-t)^N(-1).$$

At some z, $F'(z)=0$ so

$$\begin{aligned}0&=\frac{f^{(N+1)}(z)}{N!}(x-z)^N+B(N+1)(x-z)^N(-1)\\ B(N+1)(x-z)^N&=\frac{f^{(N+1)}(z)}{N!}(x-z)^N\\ B&=\frac{f^{(N+1)}(z)}{(N+1)!}.\end{aligned}$$

Now we can write

$$F(t)=\sum_{n=0}^{N}\frac{f^{(n)}(t)}{n!}(x-t)^n+\frac{f^{(N+1)}(z)}{(N+1)!}(x-t)^{N+1}.$$

Recalling that $F(a)=f(x)$ we get

$$f(x)=F(a)=\sum_{n=0}^{N}\frac{f^{(n)}(a)}{n!}(x-a)^n+\frac{f^{(N+1)}(z)}{(N+1)!}(x-a)^{N+1},$$

which is what we wanted to show. ♣

Note:

1. In essence, Taylor's Theorem says that

$$f(x)=T_n(x)+R_n(x),$$

where T_n is the n-th degree Taylor polynomial and R_n is the so-called **remainder** term.

2. We often estimate the remainder

$$R_n(x)=\frac{f^{(n+1)}(z)(x-a)^{n+1}}{(n+1)!}$$

without knowing the value of z as will be seen in Example 6.80.

3. An important consequence of Taylor's Theorem is that if $\lim\limits_{n\to\infty}R_n(x)=0$ for all x in the open interval I, then

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x-a)^n,$$

i.e. the Taylor series centred at a of f converges to f on I.

Example 6.80: Approximating Sine

Find a polynomial approximation for $\sin x$ *accurate to* ± 0.005 *for values of* x *in* $[-\pi/2, \pi/2]$.

Solution. From Taylor's Theorem with $a = 0$:

$$\sin x = \sum_{n=0}^{N} \frac{f^{(n)}(0)}{n!} x^n + \frac{f^{(N+1)}(z)}{(N+1)!} x^{N+1}.$$

What can we say about the size of the term

$$\frac{f^{(N+1)}(z)}{(N+1)!} x^{N+1}?$$

Every derivative of $\sin x$ is $\pm \sin x$ or $\pm \cos x$, so $|f^{(N+1)}(z)| \le 1$.

So we need to pick N so that

$$\left| \frac{x^{N+1}}{(N+1)!} \right| < 0.005.$$

Since we have limited x to $[-\pi/2, \pi/2]$,

$$\left| \frac{x^{N+1}}{(N+1)!} \right| < \left(\frac{\pi}{2}\right)^N + \frac{1}{N+1}! < \frac{2^{N+1}}{(N+1)!}.$$

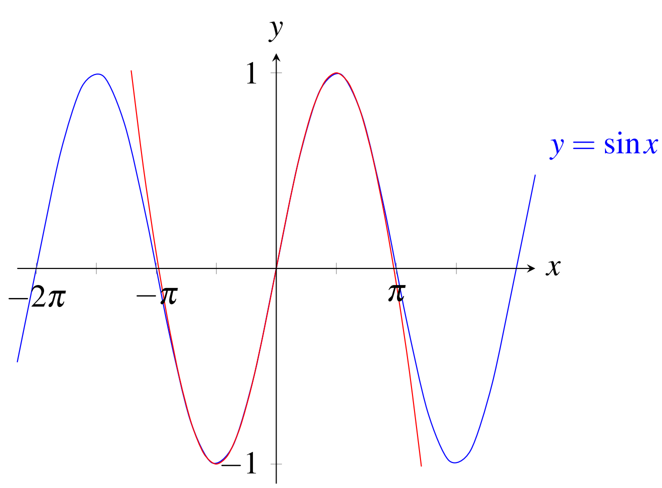
The quantity on the right decreases with increasing N, so all we need to do is find an N so that

$$\frac{2^{N+1}}{(N+1)!} < 0.005.$$

A little trial and error shows that $N = 8$ works, and in fact $2^9/9! < 0.0015$, so

$$\begin{aligned}\sin x &= \sum_{n=0}^{8} \frac{f^{(n)}(0)}{n!} x^n \pm 0.0015 \\ &= x - \frac{x^3}{6} + \frac{x^5}{120} - \frac{x^7}{5040} \pm 0.0015.\end{aligned}$$

The graphs of $\sin x$ and and the approximation are shown below. As x gets larger, the approximation heads to negative infinity very quickly, since it is essentially acting like $-x^7$.



♣

Note: We can now approximate the value of $\sin(x)$ to within 0.005 by using simple trigonometric identities to translate x into the interval $[-\pi/2, \pi/2]$.

We can extract a bit more information from this example.

Example 6.81: Convergence of Power Series Representation of Sine

Show that

$$\sin x = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$

by showing that $\lim_{n\to\infty} R_n(x) = 0$.

Solution. If we do not limit the value of x, we still have

$$\left| \frac{f^{(N+1)}(z)}{(N+1)!} x^{N+1} \right| \leq \left| \frac{x^{N+1}}{(N+1)!} \right|$$

so that $\sin x$ is represented by

$$\sum_{n=0}^{N} \frac{f^{(n)}(0)}{n!} x^n \pm \left| \frac{x^{N+1}}{(N+1)!} \right|.$$

If we can show that

$$\lim_{N\to\infty} \left| \frac{x^{N+1}}{(N+1)!} \right| = 0$$

for each x then

$$\sin x = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!},$$

that is, the sine function is actually equal to its Maclaurin series for all x. How can we prove that the limit is zero? Suppose that N is larger than $|x|$, and let M be the largest integer less than $|x|$ (if $M = 0$ the following is even easier). Then

$$\begin{aligned}
\frac{|x^{N+1}|}{(N+1)!} &= \frac{|x|}{N+1} \frac{|x|}{N} \frac{|x|}{N-1} \cdots \frac{|x|}{M+1} \frac{|x|}{M} \frac{|x|}{M-1} \cdots \frac{|x|}{2} \frac{|x|}{1} \\
&\leq \frac{|x|}{N+1} \cdot 1 \cdot 1 \cdots 1 \cdot \frac{|x|}{M} \frac{|x|}{M-1} \cdots \frac{|x|}{2} \frac{|x|}{1} \\
&= \frac{|x|}{N+1} \frac{|x|^M}{M!}.
\end{aligned}$$

The quantity $|x|^M/M!$ is a constant, so

$$\lim_{N\to\infty} \frac{|x|}{N+1} \frac{|x|^M}{M!} = 0$$

and by the Squeeze Theorem 6.7

$$\lim_{N\to\infty} \left| \frac{x^{N+1}}{(N+1)!} \right| = 0$$

as desired. ♣

Essentially the same argument works for $\cos x$ and e^x. Unfortunately, it is more difficult to show that most functions are equal to their Maclaurin series.

Example 6.82: Approximating e

Find a polynomial approximation for e^x near $x = 2$ accurate to ± 0.005.

Solution. From Taylor's Theorem:

$$e^x = \sum_{n=0}^{N} \frac{e^2}{n!}(x-2)^n + \frac{e^z}{(N+1)!}(x-2)^{N+1},$$

since $f^{(n)}(x) = e^x$ for all n. We are interested in x near 2, and we need to keep $|(x-2)^{N+1}|$ in check, so we may as well specify that $|x-2| \leq 1$, so $x \in [1,3]$. Also

$$\left| \frac{e^z}{(N+1)!} \right| \leq \frac{e^3}{(N+1)!},$$

so we need to find an N that makes $e^3/(N+1)! \leq 0.005$. This time $N = 6$ makes $e^3/(N+1)! < 0.004$, so the approximating polynomial is

$$e^x = e^2 + e^2(x-2) + \frac{e^2}{2}(x-2)^2 + \frac{e^2}{6}(x-2)^3 + \frac{e^2}{24}(x-2)^4 + \frac{e^2}{120}(x-2)^5 \pm 0.004.$$

Note that our approximation requires that we already have a very accurate approximation of the value e^2, which we shouldn't assume we have in the context of trying to approximate e^x. For this reason we typically try to centre our series on values for which the derivative of the function is easy to evaluate (e.g. $a = 0$). ♣

Note well that in these examples we found polynomials of a certain accuracy only on a small interval, even though the series for $\sin x$ and e^x converge for all x. This is typical. To get the same accuracy on a larger interval would require more terms.

Exercises for Section 6.8.5

Exercise 6.8.11 *Find a polynomial approximation for each of the functions on the given interval within the stated error.*

(a) $f(x) = \cos x$, $[0, \pi]$, $\pm 10^{-3}$

(b) $f(x) = \ln x$, $[1/2, 3/2]$, 10^{-3}

(c) $f(x) = \ln x$, $[1, 3/2]$, 10^{-3}

Exercise 6.8.12 *Show that each function is equal to its Taylor series for all x by showing that* $\lim_{n\to\infty} R_n(x) = 0$.

(a) $f(x) = \cos x$

(b) $f(x) = e^x$

Selected Exercise Answers

1.2.1 10

1.2.2 $35/3$

1.3.1 x^2

1.3.2 $2x^2$

1.3.3 $2x^2 - 8$

1.3.4 $2b^2 - 2a^2$

1.3.5 4 rectangles: $41/4 = 10.25$, 8 rectangles: $183/16 = 11.4375$

1.3.6 $23/4$

1.4.1
(a) 8
(b) 2
(c) 9.3339
(d) 2

1.4.2
(a) $87/2$
(b) 2
(c) $\ln(10)$
(d) $e^5 - 1$
(e) $3^4/4$
(f) $2^6/6 - 1/6$
(g) $16\sqrt{2} - 2$
(h) $\dfrac{2}{3}$
(i) $\dfrac{2\sqrt{2}}{3}$
(j) $\dfrac{\log 14 - \log 5}{3}$
(k) -1

1.4.3
(a) $x^2 - 3x$
(b) $2x(x^4 - 3x^2)$
(c) e^{x^2}
(d) $2xe^{x^4}$
(e) $\tan(x^2)$
(f) $2x\tan(x^4) - 10\tan(100x^2)$

1.4.4 31, 14

1.4.5 5

1.4.7 (a) \$530.40

(b) \$119.60

1.4.8 (a) \$7500

(b) \$2500

1.5.1 (a) $(16/3)x^{3/2}+C$

(b) t^3+t+C

(c) $8\sqrt{y}+C$

(d) $-2/z+C$

(e) $7\ln s+C$

(f) $(5x+1)^3/15+C$

(g) $(t-6)^3/3+C$

(h) $2z^{5/2}/5+C$

(i) $-4/\sqrt{x}+C$

(j) $4t-t^2+C,\ t<2;\ t^2-4t+8+C,\ t\geq 2$

(k) $(\sin x-4)\cdot\cos x+C$

(l) $y^2-\tan y+C$

(m) $\cot x+\sec x+C$

1.5.2 (a) $\frac{s}{6}\left(4s^2+3s-6\right)+C$

(b) $s^3-0.5s^2+s-2\log s+C$

(c) $\dfrac{u^{8/3}-3\sqrt[3]{2}}{4u^{2/3}}+C$

(d) $\dfrac{8\sqrt{t}+t\log t-4}{4t}+C$

1.5.3 (a) $f(t)=\frac{3}{2}t^2+2t+5$

(b) $f(t)=t+3\log t-2$

(c) $f(t)=-\cos t-1$

(d) $f(t)=0.5e^{2t}-4t$

1.5.4 (a) $R(q)=16q-0.004q^2$

(b) $p(q)=\frac{16q-0.004q^2}{q}$

1.5.5 \$36,400 when $q=5000$ heads.

1.5.6 $C(q)=0.0005q^2+50q+500$

1.5.7 (a) $N(t)=\frac{-2}{3}t^3+5t^2+100t+1$ for $0\leq t\leq 5$.

(b) 272

2.1.1

$-(1-t)^{10}/10+C$

(b) $x^5/5+2x^3/3+x+C$

(c) $(x^2+1)^{101}/202+C$

(d) $-3(1-5t)^{2/3}/10+C$

(e) $(\sin^4 x)/4+C$

(f) $-(100-x^2)^{3/2}/3+C$

(g) $-2\sqrt{1-x^3}/3+C$

(h) $\sin(\sin \pi t)/\pi+C$

(i) $1/(2\cos^2 x)=(1/2)\sec^2 x+C$

(j) $-\ln|\cos x|+C$

(k) $\tan^2(x)/2+C$

(l) $-\cos(\tan x)+C$

(m) $(27/8)(x^2-7)^{8/9}$

(n) $f(x)^2/2$

2.1.2 (a) 0

(b) $1/4$

(c) $1/10$

(d) $\sqrt{3}/4$

(e) $-(3^7+1)/14$

(f) 0

2.1.3 $15e^{-0.2t}+9985$ trucks.

2.1.4 $p(q)=100(1/(1-q)+\log(q-1))+102$

2.1.5 $p(q)=\frac{200}{\sqrt{q^2+3}}-97$

2.1.2 Approximately 300 units.

2.2.1 (a) $x/2-\sin(2x)/4+C$

(b) $-\cos x+(\cos^3 x)/3+C$

(c) $3x/8-(\sin 2x)/4+(\sin 4x)/32+C$

(d) $(\cos^5 x)/5-(\cos^3 x)/3+C$

(e) $\sin x-(\sin^3 x)/3+C$

(f) $(\sin^3 x)/3-(\sin^5 x)/5+C$

(g) $-2(\cos x)^{5/2}/5+C$

(h) $\tan x-\cot x+C$

(i) $(\sec^3 x)/3-\sec x+C$

(j) $-\cos x+\sin x+C$

(k) $\dfrac{3}{2}\ln|\sec x+\tan x|+\tan x+\dfrac{1}{2}\sec x\tan x+C$

(l) $\dfrac{\tan^5(x^2)}{10}+C$

(m) $\dfrac{\tan^5(x^2)}{10}+C$

2.3.1 (a) $x\sqrt{x^2-1}/2-\ln|x+\sqrt{x^2-1}|/2+C$

(b) $x\sqrt{9+4x^2}/2+(9/4)\ln|2x+\sqrt{9+4x^2}|+C$

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

(d) $\arcsin(x)/8-\sin(4\arcsin x)/32+C$

(e) $\ln|x+\sqrt{1+x^2}|+C$

(f) $(x+1)\sqrt{x^2+2x}/2-\ln|x+1+\sqrt{x^2+2x}|/2+C$

(g) $-\arctan x-1/x+C$

(h) $2\arcsin(x/2)-x\sqrt{4-x^2}/2+C$

(i) $\arcsin(\sqrt{x})-\sqrt{x}\sqrt{1-x}+C$

(j) $(2x^2+1)\sqrt{4x^2-1}/24+C$

2.4.1 (a) $(1/2)e^{x^2}+C$

(b) $(x/2)-\sin(2x)/4+C=$
$(x/2)-(\sin x\cos x)/2+C$

(c) $(x^2\arctan x+\arctan x-x)/2+C$

(d) $x^2/4-(\cos^2 x)/4-(x\sin x\cos x)/2+C$

(e) $x/4-(x\cos^2 x)/2+(\cos x\sin x)/4+C$

2.4.2 (a) $\cos x+x\sin x+C$

(b) $x^2\sin x-2\sin x+2x\cos x+C$

(c) $(x-1)e^x+C$

(d) $x\ln x-x+C$

(e) $-x^3\cos x+3x^2\sin x+6x\cos x-6\sin x+C$

(f) $x^3\sin x+3x^2\cos x-6x\sin x-6\cos x+C$

(g) $x\arctan(\sqrt{x})+\arctan(\sqrt{x})-\sqrt{x}+C$

(h) $2\sin(\sqrt{x})-2\sqrt{x}\cos(\sqrt{x})+C$

(i) $\sec x\csc x-2\cot x+C$

2.5.1 (a) $-\ln|x-2|/4+\ln|x+2|/4+C$

(b) $-x^3/3-4x-4\ln|x-2|+$
$4\ln|x+2|+C$

(c) $-1/(x+5)+C$

(d) $-x-\ln|x-2|+\ln|x+2|+C$

(e) $-4x+x^3/3+8\arctan(x/2)+C$

(f) $(1/2)\arctan(x/2+5/2)+C$

(g) $x^2/2-2\ln(4+x^2)+C$

(h) $(1/4)\ln|x+3|-(1/4)\ln|x+7|+C$

(i) $(1/5)\ln|2x-3|-(1/5)\ln|1+x|+C$

(j) $(1/3)\ln|x|-(1/3)\ln|x+3|+C$

2.6.1 (a) T,S: 4 ± 0

(b) T: 9.28125 ± 0.281125; S: 9 ± 0

(c) T: 60.75 ± 1; S: 60 ± 0

(d) T: 1.1167 ± 0.0833; S: 1.1000 ± 0.0167

(e) T: 0.3235 ± 0.0026; S: 0.3217 ± 0.000065

(f) T: 0.6478 ± 0.0052; S: 0.6438 ± 0.000033

(g) T: 2.8833 ± 0.0834; S: 2.9000 ± 0.0167

(h) T: 1.1170 ± 0.0077; S: 1.1114 ± 0.0002

(i) T: 1.097 ± 0.0147; S: 1.089 ± 0.0003

(j) T: 3.63 ± 0.087; S: 3.62 ± 0.032

2.7.1 (a) Converges to $\pi/2$

(b) Divergent (to ∞)

(c) Converges to 1

(d) Divergent (to ∞)

(e) Converges to $\frac{5}{3}(4^{3/5})$

(f) Converges to 1

(g) Divergent

(h) Converges to 1/3

(i) Divergent

(j) Converges to $\frac{\pi^2}{2+2\pi^2}$

(k) Converges to π

(l) Divergent

2.7.3 $0 < p < 1$

2.8.1 $\dfrac{(t+4)^4}{4} + C$

2.8.2 $\dfrac{(t^2-9)^{5/2}}{5} + C$

2.8.3 $\dfrac{(e^{t^2}+16)^2}{4} + C$

2.8.4 $\cos t - \dfrac{2}{3}\cos^3 t + C$

2.8.5 $\dfrac{\tan^2 t}{2} + C$

2.8.6 $\ln|t^2+t+3| + C$

2.8.7 $\dfrac{1}{8}\ln|1-4/t^2| + C$

2.8.8 $\dfrac{1}{25}\tan(\arcsin(t/5)) + C = \dfrac{t}{25\sqrt{25-t^2}} + C$

2.8.9 $\dfrac{2}{3}\sqrt{\sin 3t} + C$

2.8.10 $t\tan t + \ln|\cos t| + C$

2.8.11 $2\sqrt{e^t+1} + C$

2.8.12 $\dfrac{3t}{8} + \dfrac{\sin 2t}{4} + \dfrac{\sin 4t}{32} + C$

2.8.13 $\dfrac{\ln|t|}{3} - \dfrac{\ln|t+3|}{3} + C$

2.8.14 $\dfrac{-1}{\sin\arctan t} + C = -\sqrt{1+t^2}/t + C$

2.8.15 $\dfrac{-1}{2(1+\tan t)^2} + C$

2.8.16 $\dfrac{(t^2+1)^{5/2}}{5} - \dfrac{(t^2+1)^{3/2}}{3} + C$

2.8.17 $\dfrac{e^t \sin t - e^t \cos t}{2} + C$

2.8.18 $\dfrac{(t^{3/2} + 47)^4}{6} + C$

2.8.19 $\dfrac{2}{3(2-t^2)^{3/2}} - \dfrac{1}{(2-t^2)^{1/2}} + C$

2.8.20 $\dfrac{\ln|\sin(\arctan(2t/3))|}{9} + C = (\ln(4t^2) - \ln(9+4t^2))/18 + C$

2.8.21 $\dfrac{(\arctan(2t))^2}{4} + C$

2.8.22 $\dfrac{3\ln|t+3|}{4} + \dfrac{\ln|t-1|}{4} + C$

2.8.23 $\dfrac{\cos^7 t}{7} - \dfrac{\cos^5 t}{5} + C$

2.8.24 $\dfrac{-1}{t-3} + C$

2.8.25 $\dfrac{-1}{\ln t} + C$

2.8.26 $\dfrac{t^2(\ln t)^2}{2} - \dfrac{t^2 \ln t}{2} + \dfrac{t^2}{4} + C$

2.8.27 $(t^3 - 3t^2 + 6t - 6)e^t + C$

2.8.28 $\dfrac{5+\sqrt{5}}{10}\ln(2t+1-\sqrt{5}) + \dfrac{5-\sqrt{5}}{10}\ln(2t+1+\sqrt{5}) + C$

3.1.1 (a) $2/\pi$; $2/\pi$; 0

(b) $4/3$

(c) $1/a$

(d) $\pi/4$

3.1.2 (a) $8\sqrt{2}/15$

(b) $2/3 - 2/\pi$

(c) $3/\pi - 3\sqrt{3}/(2\pi) - 1/8$

(d) $1/3$

(e) $1/5$

(f) $1/6$

3.1.3 (a) $1/12$

(b) $500/3$

3.1.4 (a) $9/2$

(b) $4/3$

(c) $10\sqrt{5}/3-6$

(d) 2

3.1.5 (a) 6.75

(b) 12.15

(c) 16/3

(d) $3\log(2)$

(e) 4.5

(f) 1/3

3.1.6 It rises until $t=100/49$, then falls. The position of the object at time t is $s(t)=-4.9t^2+20t+k$. The net distance travelled is $-45/2$, that is, it ends up $45/2$ meters below where it started. The total distance travelled is $6205/98$ meters.

3.1.7 $\displaystyle\int_0^{2\pi} \sin t\,dt = 0$

3.1.8 net: 2π, total: $2\pi/3+4\sqrt{3}$

3.1.9 8

3.1.10 $17/3$

3.1.11 $A=18$, $B=44/3$, $C=10/3$

3.3.1 (a) $(1/3)(\text{area of base})(\text{height})$

(b) 2π

3.3.2 (a) $4\pi/3$

(b) $50\pi/3$

(c) $\pi/2$

(d) $\pi/4$

3.3.3 (a) $26\pi/5$

(b) $\pi/7$

(c) $16\pi/3$

(d) π

(e) 1.5420

(f) $8\pi/3$

(g) $8\pi/3$

(h) $\pi(\pi/2-1)$

3.3.4 (a) $2\pi/4$

(b) $256\sqrt{2}\pi/13$

(c) π

(d) $\pi/2$

3.3.6 (a) $16\pi/3$

(b) $9\pi/2$

(c) $117\pi/5$

(d) $625\pi/3$

3.3.7 16π, 24π

3.3.9 $\pi h^2(3r-h)/3$

3.4.5 (a) $114\pi/5$

(b) $74\pi/5$

(c) 20π

(d) 4π

4.1.1 (a) $z=y^2$, $z=x^2$, $z=0$, lines of slope 1

(b) $z=|y|$, $z=|x|$, $z=2|x|$, diamonds

(c) $z=e^{-y^2}\sin(y^2)$, $z=e^{-x^2}\sin(x^2)$, $z=e^{-2x^2}\sin(2x^2)$, circles

(d) $z = -\sin(y)$, $z = \sin(x)$, $z = 0$, lines of slope 1

(e) $z = y^4$, $z = x^4$, $z = 0$, hyperbolas

4.1.2 (a) $\{(x,y) \mid |x| \leq 3 \text{ and } |y| \geq 2\}$

(b) $\{(u,v) \mid u^2 + v^2 \leq 2\}$

(c) $\{(x,y) \mid x^2 + 4y^2 \leq 16\}$

(d) All real x and y.

(e) All real x,y, and z

(f) $\{(x,y) \mid 1 \leq x^2 + y^2 \leq 3\}$

(g) $\{(u,v) \mid u \neq v\}$

(h) All real s and t

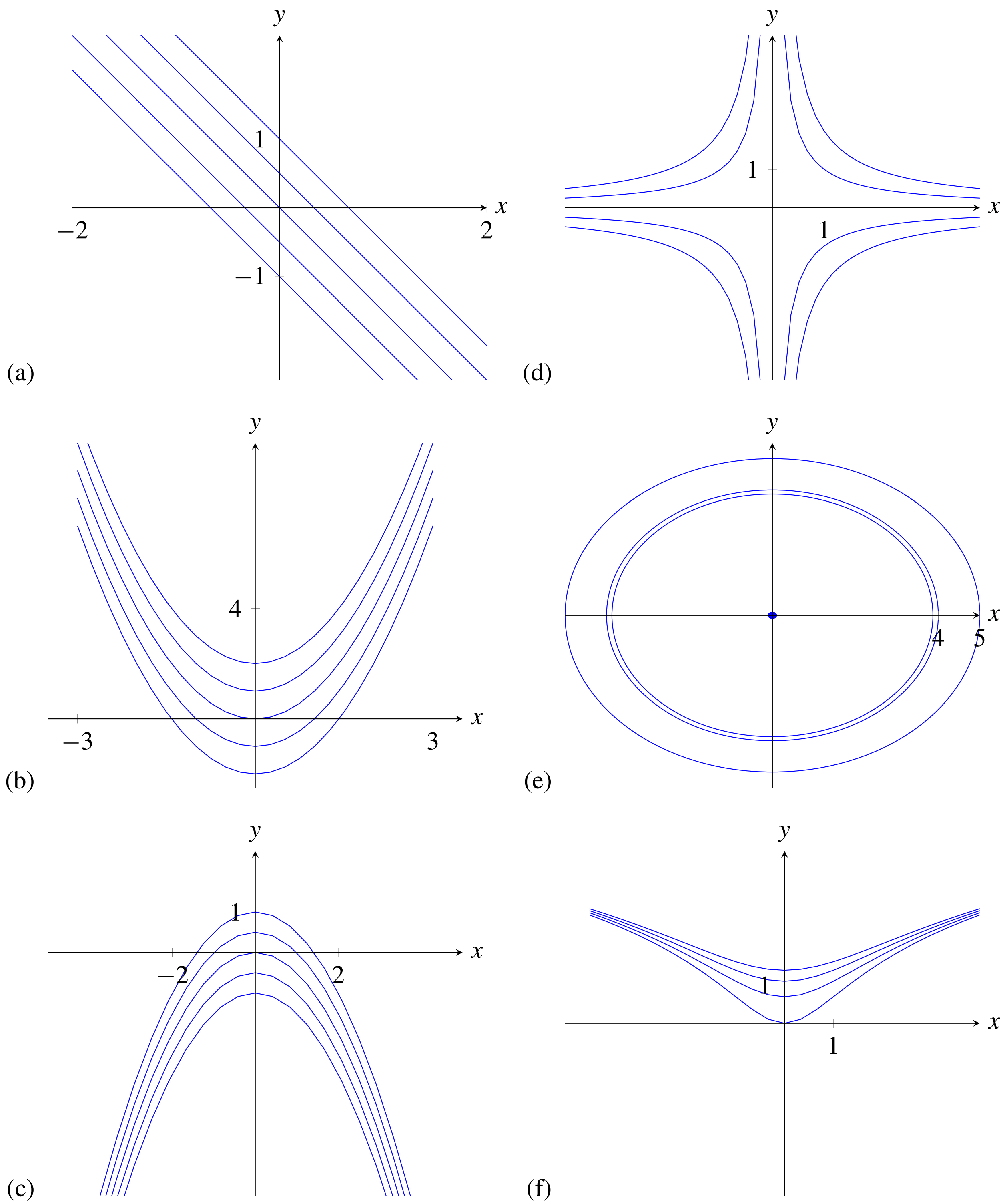
(i) $\{(r,s) \mid rs > 0\}$

(j) All real x and y

(k) $\{(x,y) \mid x + y > 7\}$

4.1.4

We draw the level curves of f:



4.1.5 (a) $R(q_c, q_b) = \frac{1}{420}\left(-105q_c^2 - 102q_cq_b + 42000q_c - 140q_b^2 + 63000q_b\right)$

(b) The region where all of $p_c, p_b, q_c, q_b \geq 0$.

4.1.6 (a) $R(x,y) = -0.1q_b^2 - 0.9q_bq_c + 10q_b - q_c^2 + 30q_c$

(b) The region where all of $p_b, p_c, q_b, q_c \geq 0$.

4.1.7 \$50,148.53; \$28,661.05

4.1.8 About 39 tires.

4.2.1 (a) 16

(b) 4

(c) $15/8$

(d) $1/2$

(e) $5/6$

(f) $12-65/(2e)$

(g) $1/2$

(h) $\pi/64$

(i) $(2/9)2^{3/2}-(2/9)$

(j) $(1-\cos(1))/4$

(k) $(2\sqrt{2}-1)/6$

(l) $\pi-2$

(m) $(e^9-1)/6$

(n) $\dfrac{4}{15}-\dfrac{\pi}{4}$

(o) $1/3$

4.2.2 (a) $4/5$

(b) 8π

(c) 2

(d) $5/3$

(e) $81/2$

(f) $2a^3/3$

(g) 4π

(h) $\pi/32$

(i) $31/8$

(j) $128/15$

4.2.3 448

4.2.4 1800π m^3

4.2.5 $\dfrac{(e^2+8e+16)}{15}\sqrt{e+4}-\dfrac{5\sqrt{5}}{3}-\dfrac{e^{5/2}}{15}+\dfrac{1}{15}$

4.3.1 (a) $11/24$

(b) $623/60$

(c) $-3e^2/4+2e-3/4$

(d) $1/20$

(e) $\pi/48$

(f) $11/84$

(g) $151/60$

(h) π

4.3.2 $\dfrac{3\pi}{16}$

4.4.2 (a) $f(x)=\frac{2}{9}(3-x), \quad x\in[0,3]$

(b) $f(x)=\frac{4}{x^5}, \quad x\in[1,\infty)$

(c) $f(x,y)=6xy^2, \quad (x,y)\in[0,1]\times[0,1]$

(d) $f(x,y)=\frac{3e}{7}x^2e^{-y}, \quad (x,y)\in[1,2]\times[1,\infty)$

4.4.3 (a) i.0, ii. 0.173, iii.$\approx$ 1, iv. 0.0498

(b) i. 0.739, ii. 0.261, iii.0.599, iv. 0.739

(c) i. 0.25, ii. 0

(d) i. 0.0469, ii. 0.109

4.4.5 No; $f(-1) < 0$

4.4.6 (a) 0.318

(b) 0.0067

4.4.7 (c) i.3, ii.3, iii. $\sqrt{3}$

4.4.8 (a) 0.2

(b) 0.5

4.4.9 2 minutes

4.4.10 5 kg

4.4.11 (a) Not independent

4.4.12 (a) i.3/4,4/7; ii.3/80,19/392; iii.1/63

5.1.1 (a) third order, non-linear

(b) third order, linear

(c) first order, non-linear

(d) fourth order, non-linear

(e) second order, linear

(f) first order, non-linear

5.1.2 (a) non-homogeneous

(b) homogeneous

(c) non-homogeneous

(d) homogeneous

(e) non-homogeneous

(f) homogeneous

5.1.3 (a) has constant coefficients

(b) does not have constant coefficients

(c) does not have constant coefficients

(d) does not have constant coefficients

(e) does not have constant coefficients

(f) does not have constant coefficients

5.2.2 (a) $y = n\pi$, for any integer n.

(b) none

5.2.3 (a) $y = \arctan t + C$

(b) $y = t\ln t - t + C$

(c) $y = \pm\sqrt{t^2 + C}$

(d) $y = \pm 1$, $y = (1 + Ae^{2t})/(1 - Ae^{2t})$

(e) $y^4/4 - 5y = t^2/2 + C$

(f) $y = M + Ae^{-kt}$

5.2.4 (a) $y = \dfrac{t^{n+1}}{n+1} + 1$

(b) $y = (2t/3)^{3/2}$ and $y(t) = 0$

(c) $y = 2e^{3t/2}$

5.2.5 $\dfrac{10\ln(15/2)}{\ln 5} \approx 2.52$ minutes

5.2.6 $t = -\dfrac{\ln 2}{k}$

5.2.7 $600e^{-6\ln 2/5} \approx 261$ mg; $\dfrac{5\ln 300}{\ln 2} \approx 41$ days

5.2.8 $100e^{-200\ln 2/191} \approx 48$ mg; $\dfrac{5730\ln 50}{\ln 2} \approx 32339$ years

5.2.9 $y = y_0 e^{t\ln 2}$

5.2.10 $500e^{-5\ln 2/4} \approx 210$ g

5.2.11 $y = \dfrac{M}{1 + Ae^{-Mkt}}$

5.3.1 (a) $y = Ae^{-5t}$

(b) $y = Ae^{2t}$

(c) $y = Ae^{-\arctan t}$

(d) $y = Ae^{-t^3/3}$

5.3.2 (a) $y = 4e^{-t}$

(b) $y = -2e^{3t-3}$

(c) $y = e^{1+\cos t}$

(d) $y = e^2 e^{-e^t}$

(e) $y = 0$

(f) $y = 0$

(g) $y = 4t^2$

(h) $y = -2e^{(1/t)-1}$

(i) $y = e^{1-t^{-2}}$

(j) $y = 0$

5.3.3 (a) $y = Ae^{-4t} + 2$

(b) $y = Ae^{2t} - 3$

(c) $y = Ae^{-(1/2)t^2} + 5$

(d) $y = Ae^{-e^t} - 2$

(e) $y = Ae^t - t^2 - 2t - 2$

(f) $y = Ae^{-t/2} + t - 2$

5.3.4 (a) $y = At^2 - \dfrac{1}{3t}$

(b) $y = \dfrac{c}{t} + \dfrac{2}{3}\sqrt{t}$

(c) $y = A\cos t + \sin t$

(d) $y = \dfrac{A}{\sec t + \tan t} + 1 - \dfrac{t}{\sec t + \tan t}$

5.3.5 (a) $y = 2 - e^{-4t}$

(b) $y = 6e^{2t} - 3$

(c) $y = 5 - 4e^{2-(1/2)t^2}$

(d) $y = e^{-e^t} - 2$

(e) $y = 6e^t - t^2 - 2t - 2$

(f) $y = t - 2$

5.3.6 $k = \ln 5$, $y = 100e^{-t\ln 5}$

5.3.7 $k = -12/13$, $y = \exp(-13t^{1/13})$

5.3.8 $y = 10^6 e^{t\ln(3/2)}$

5.3.9 $y = 10e^{-t\ln(2)/6}$

5.4.1 (a) B

(b) A

(c) C

(d) D

5.4.4 (a) $y(1) \approx 1.355$

(b) $y(1) \approx 40.31$

(c) $y(1) \approx 1.05$

(d) $y(1) \approx 2.30$

6.1.2 1

6.1.4 (a) diverges

(b) converges to 1

(c) diverges

(d) converges to 0

(e) converges to 0

(f) converges to 1

(g) converges to 1

(h) converges to 0

6.2.1 (a) $\lim_{n\to\infty} n^2/(2n^2+1) = 1/2$

(b) $\lim_{n\to\infty} 5/(2^{1/n}+14) = 1/3$

(c) $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges, so $\sum_{n=1}^{\infty} 3\frac{1}{n}$ diverges

6.2.2 (a) $-3/2$

(b) 11

(c) 20

(d) $3/4$

(e) $3/2$

(f) $3/10$

6.3.1 (a) diverges

(b) diverges

(c) converges

(d) converges

(e) converges

(f) converges

(g) diverges

(h) converges

6.3.2 (a) $N = 5$

(b) $N = 10$

(c) $N = 1687$

(d) any integer greater than e^{200}

6.4.1 (a) converges

(b) converges

(c) diverges

(d) converges

6.4.2 (a) 0.90

(b) 0.95

6.5.1 (a) converges

(b) converges

(c) converges

(d) diverges

(e) diverges

(f) diverges

(g) converges

(h) diverges

(i) converges

(j) diverges

6.6.1 (a) converges absolutely

(b) diverges

(c) converges conditionally

(d) converges absolutely

(e) converges conditionally

(f) converges absolutely

(g) diverges

(h) converges conditionally

6.7.2 (a) converges

(b) converges

(c) converges

(d) diverges

6.8.1 (a) $R=1, I=(-1,1)$

(b) $R=\infty, I=(-\infty,\infty)$

(c) $R = e$, $I = (2-e, 2+e)$

(d) $R = 0$, converges only when $x = 2$

(e) $R = 1$, $I = [-6, -4]$

(f) $R = e$

6.8.2 (a) the alternating harmonic series

6.8.3 (a) $\sum_{n=0}^{\infty}(n+1)x^n$

(b) $\sum_{n=0}^{\infty}(n+1)(n+2)x^n$

(c) $\sum_{n=0}^{\infty}\frac{(n+1)(n+2)}{2}x^n$, $R = 1$

(d) $C + \sum_{n=0}^{\infty}\frac{-1}{(n+1)(n+2)}x^{n+2}$

6.8.4 (a) $\sum_{n=0}^{\infty}(-1)^n x^{2n}/(2n)!$, $R = \infty$

(b) $\sum_{n=0}^{\infty}x^n/n!$, $R = \infty$

6.8.5 (a) $\sum_{n=0}^{\infty}(-1)^n\frac{(x-5)^n}{5^{n+1}}$, $R = 5$

(b) $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{(x-1)^n}{n}$, $R = 1$

(c) $\ln(2) + \sum_{n=1}^{\infty}(-1)^{n-1}\frac{(x-2)^n}{n2^n}$, $R = 2$

(d) $\sum_{n=0}^{\infty}(-1)^n(n+1)(x-1)^n$, $R = 1$

6.8.6 (a) $1 + \sum_{n=1}^{\infty}\frac{1\cdot 3\cdot 5\cdots(2n-1)}{n!2^n}x^n = 1 + \sum_{n=1}^{\infty}\frac{(2n-1)!}{2^{2n-1}(n-1)!\,n!}x^n$, $R = 1$

6.8.7 $T_5(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!}$

(a) $\sin(0.1) \approx T_5(0.1) \approx 0.10016675$

(b) $\sin(0.1) = 0.0998334\ldots$ using a calculator. Our approximation is accurate to $0.10016675 - 0.0998334\ldots = 0.000\bar{3}$.

6.8.8 $T_3(x) = x + x^2 + x^3$. The point $x = 5$ is not close to $x = 0$, and f is not continuous at $x = 1$.

6.8.9 (a) $f^{(n)}(x) = \frac{(-1)^{(n-1)}(n-1)!}{x^n}$

(b) $T_n(x) = \ln(1) + \displaystyle\sum_{i=1}^{n} \frac{\left(\frac{(-1)^{(i-1)}(i-1)!}{1^n}\right)}{i!}(x-1)^i = \sum_{i=1}^{n} \left(\frac{(-1)^{(i-1)}(i-1)!}{i!}\right)(x-1)^i$ since $\ln(1) = 0$ and $1^n = 1$.

6.8.10 0.262364...

6.8.11 (a) $1 - \dfrac{x^2}{2} + \dfrac{x^4}{24} - \dfrac{x^6}{720} + \cdots + \dfrac{x^{12}}{12!}$

Index